



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE ESTUDIOS SUPERIORES IZTACALA

Análisis *in silico* de los genes de RNA de
transferencia y de RNA ribosomal 5S en
Leishmania major

T E S I S

QUE PARA OBTENER EL GRADO ACADÉMICO DE

**MAESTRA EN CIENCIAS BIOLÓGICAS
(BIOLOGÍA EXPERIMENTAL)**

P R E S E N T A

NORMA EDITH PADILLA MEJÍA

DIRECTOR DE TESIS: DR. SANTIAGO MARTÍNEZ CALVILLO

TLALNEPANTLA, EDO. MEX.

MAYO, 2009



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

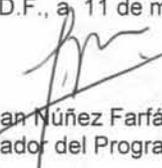
Dr. Isidro Ávila Martínez
Director General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día 9 de febrero de 2009, se aprobó el siguiente jurado para el examen de grado de **MAESTRA EN CIENCIAS BIOLÓGICAS (BIOLOGÍA EXPERIMENTAL)** de la alumna **NORMA EDITH PADILLA MEJÍA** con número de cuenta **9639167-3** con la tesis titulada "**Análisis *in silico* de los genes de RNA de transferencia y de RNA ribosomal 5S en *Leishmania major***" realizada bajo la dirección del **DR. SANTIAGO MARTÍNEZ CALVILLO**:

Presidente: DR. ROBERTO HERNÁNDEZ FERNÁNDEZ
Vocal: DRA. ANA MARÍA CEVALLOS GAOS
Secretario: DR. SANTIAGO MARTÍNEZ CALVILLO
Suplente: DRA. MARÍA IMELDA LÓPEZ VILLASEÑOR
Suplente: DRA. ROSAURA HERNÁNDEZ RIVAS

Sin otro particular, me es grato enviarle un cordial saludo.

Atentamente
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, D.F., a 11 de mayo de 2009.


Dr. Juan Núñez Farfán
Coordinador del Programa

c.c.p. Expediente del (la) interesado (a).

AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México, que ha sido mi *Alma Mater* durante toda mi formación profesional y me ha brindado muchas oportunidades.

Al Posgrado en Ciencias Biológicas de la UNAM

Agradezco igualmente el apoyo de Becas Nacionales de Posgrado recibido por parte de CONACYT (becario 216026).

A mi tutor principal, el Dr. Santiago Martínez Calvillo, por permitirme ser parte de su grupo de trabajo y por todas sus enseñanzas profesionales y personales.

A los miembros de mi Comité Tutorial

A la Dra. Ana María Cevallos Gaos

A la Dra. Rosaura Hernández Rivas

por sus valiosos comentarios que sirvieron para enriquecer este trabajo, también les agradezco su apoyo y aliento durante el proyecto

A los integrantes del jurado de examen de grado:

Dr. Santiago Martínez Calvillo, Dra. Ana María Cevallos Gaos, Dr. Roberto Hernández Fernández, Dra. Rosaura Hernández Rivas y Dra. María Imelda López Villaseñor.

Les agradezco su ayuda en la revisión de la tesis y sus comentarios para enriquecer el presente trabajo.

DEDICATORIAS

Con todo mi amor para

Mi mamá Lucía Mejía Lazcano

Mi papá Carlos M. Padilla Rodríguez

Gracias por su apoyo incondicional, por su aliento y su ayuda en todos mis proyectos e inquietudes y siempre querer lo mejor para mí y mis hermanos. Y porque todo lo que me enseñaron me permite ahora alcanzar mis metas, este trabajo es para ustedes.

Y para mis hermanos

Carlos G. y Sandra Paola Padilla Mejía

Con mucho cariño y amor. Se que cumplirán todos sus metas y sueños y que los que han logrado hasta ahora son sólo el comienzo. Recuerden que siempre podrán contar conmigo. Les dedico esta tesis.

ÍNDICE

Lista de Abreviaturas	III
Lista de Tablas	III
Lista de Figuras	IV
RESUMEN	1
ABSTRACT	3
1. INTRODUCCIÓN	5
1.1. Generalidades sobre tripanosomátidos	5
1.1.1. Leishmaniasis: importancia médica, epidemiología, manifestaciones clínicas	5
1.1.1.1. Ciclo de vida de <i>Leishmania</i>	7
1.1.2. Aspectos generales de <i>Trypanosoma brucei</i>	7
1.1.3. Aspectos generales de <i>Trypanosoma cruzi</i>	9
1.2. Particularidades biológicas y organización genómica de tripanosomátidos	10
1.3. Generalidades sobre la transcripción por RNA polimerasa III (Pol III)	11
1.3.1. Elementos Promotores de la Transcripción por Pol III	12
1.3.2. Factores de transcripción asociados a la RNA polimerasa III	14
1.3.3. Terminación de la transcripción por Pol III	15
1.4. Generalidades del RNA ribosomal 5S	16
1.4.1. Función del RNA ribosomal 5S	16
1.4.2. Estructura secundaria del RNAr 5S	16
1.4.3. Distribución en el genoma y expresión del RNAr 5S	17
1.5. Generalidades sobre RNAs de transferencia	18
1.5.1. Estructura y función de los RNAs de transferencia	18
1.5.2. RNAt isoaceptores	20
1.5.3. Uso de codones	22
1.5.4. RNAts isodecodificadores	23
1.5.5. RNAt-Met de iniciación y de elongación	24
1.5.6. Distribución genómica	25
1.6. RNAt Selenocisteína	26
1.6.1. Estructura y función del RNAt-Sec	27
1.6.2. Transcripción y promotores del RNAt-Sec	28
2. ANTECEDENTES	30

3. OBJETIVOS	31
3.1. Objetivo General	31
3.2. Objetivos Específicos	31
4. MÉTODOS	32
5. RESULTADOS	34
5.1. Número de genes de RNA transferencia en tripanosomátidos	34
5.2. RNAt isoceptores	41
5.3. RNAt de Metionina elongador e iniciador	44
5.4. Organización de los genes de RNAt	47
5.5. Relación espacial entre los genes de Pol II y Pol III en tripanosomátidos	49
5.6. Sintenia entre genes de Pol III	50
5.7. Intrones en los genes de RNAt en tripanosomátidos	52
5.8. Secuencias consenso de los elementos promotores	53
5.9. Señales de término de la transcripción en genes de RNAt	57
5.10. RNAt isodecodificadores	60
5.11. Uso de codones en tripanosomátidos	64
5.12. Número de genes y estructura del RNAr 5S en tripanosomátidos	66
6. DISCUSIÓN	69
7. CONCLUSIONES	87
8. BIBLIOGRAFÍA	89
9. APÉNDICES	94
Apéndice I: Datos completos sobre la localización de los genes de RNAt en Tripanosomátidos	94
Apéndice II: Señales de terminación de la transcripción de genes de RNAt en Tritryps	101

LISTA DE ABREVIATURAS

3'-UTR	Región 3' no traducida (<i>untranslated region</i>)	nt	nucleótido
AE	Elemento activador (<i>activator element</i>)	pb	pares de bases
Ala	alanina	Phe	fenilalanina
Arg	arginina	Pol	RNA polimerasa
Asn	asparagina	Pro	prolina
Asp	aspartato	PSE	Elemento proximal (<i>Proximal sequence element</i>)
Chr	cromosoma	RNAr	RNA ribosomal
Cys	cisteína	RNA _{sn}	RNA pequeño nuclear
DGC	<i>Cluster</i> de genes direccionales (<i>Directional gene cluster</i>)	RNA _t	RNA de transferencia
DSE	Secuencia distal (<i>Distal sequence element</i>)	RNA _t -e (Met-e)	RNA _t de metionina elongador (Met elongador)
EI	Elemento intermedio	RNA _t -i (Met-i)	RNA _t de metionina iniciador (Met iniciador)
Gln	glutamina	Sec	selenocisteína
Glu	glutamato	Ser	serina
Gly	glicina	SIT	Sitio de inicio de la transcripción
His	histidina	sRNA76	RNA pequeño 76 (<i>small RNA 76</i>)
Ile	isoleucina	Tb	<i>Trypanosoma brucei</i>
Leu	leucina	Tc	<i>Trypanosoma cruzi</i>
Lm	<i>Leishmania major</i>	TF	Factor de transcripción (<i>Transcription factor</i>)
Lys	lisina	Thr	treonina
max	máximo	Trp	triptofano
Met	metionina	Tyr	tirosina
min	mínimo	Val	valina

LISTA DE TABLAS

Tabla 1	Especies de <i>Leishmania</i> y la forma de leishmaniasis que producen	6
Tabla 2	Resumen de los genomas de los tripanosomátidos	11
Tabla 3	Algunos RNAs sintetizados por Pol III	11
Tabla 4	Hipótesis <i>Wobble</i> : apareamientos posibles codón:anticodón	22
Tabla 5	Principales determinantes para la función de elongador o iniciador	24
Tabla 6	Copias del RNA _t -Sec en <i>T. cruzi</i>	37
Tabla 7	Repertorio de genes de RNA _t y uso de codones en Trityps	43
Tabla 8	Lista de genes de RNA _t -Met-i y RNA _t -Met-e en Trityps	44
Tabla 9	Distribución cromosomal de genes de RNA _t en <i>L. major</i> y <i>T. brucei</i>	48
Tabla 10	Localización de los clusters de RNA _t en relación a las unidades de genes de Pol II	49
Tabla 11	Excepciones a la secuencia consenso de las cajas promotoras	
	Tabla 11- A Caja A de genes RNA _t Clase I	54
	Tabla 11- B Caja A de genes RNA _t Clase II	56
	Tabla 11- C Caja B de genes Clase I	56
Tabla 12.	Genes RNA _t isodecodificadores	63

LISTA DE FIGURAS

Figura 1	Ciclo de vida de <i>Leishmania</i> spp y manifestaciones clínicas que produce	8
Figura 2	Los tres tipos de promotores de Pol III	13
Figura 3	Diferentes vías para el reclutamiento del TFIIIB y RNA polimerasa III.	15
Figura 4	Estructura Secundaria del RNAr 5S humano	17
Figura 5	Estructura del RNAt	19
Figura 6	El código genético	21
Figura 7	Apareamientos codón : anticodón complementarios y <i>wobble</i> en eucariontes	22
Figura 8	Gen de RNAt de Arg (ACG) y el RNAt isodecodificador	23
Figura 9	Principales características entre el RNAt iniciador y el RNAt de elongación en humano	25
Figura 10	Estructura del aminoácido selenocisteína	26
Figura 11	Comparación de estructura secundaria de un RNAt convencional y el RNAt-Sec	27
Figura 12	Elementos promotores en los genes de RNAt convencionales y el RNAt-Sec	28
Figura 13	Organización de los genes de RNAt en <i>L. major</i>	35
Figura 14	Organización de los genes de RNAt en <i>T. brucei</i>	36
Figura 15	Análisis de los RNAt-Val putativos en <i>T. cruzi</i>	39
Figura 16	Análisis de Tc00.1047053507579.16, probable tRNA-Ile de <i>T. cruzi</i>	40
Figura 17	Estrategia de codificación utilizada por tripanosomátidos para poder leer los 61 codones del código genético	44
Figura 18	Estructuras secundarias del RNAt-Met iniciador y elongador en <i>L. major</i>	46
Figura 19	Sintenia de genes transcritos por Pol III en <i>Trityps</i>	51
Figura 20	Alineamiento de los genes de RNAt de tirosina en los tres tripanosomátidos	52
Figura 21	Secuencias consenso de las cajas A y B de los genes de RNAt en <i>Trityps</i>	54
Figura 22	Comparación de las cajas promotoras en los genes de RNAt-Sec	57
Figura 23	Distribución de la longitud de la señal de término de los genes de RNAt en <i>Trityps</i>	59
Figura 24	Porcentaje de genes de RNAt que presentan una segunda señal de término de la transcripción en <i>Trityps</i> y otros eucariontes	60
Figura 25	Comparación de secuencias de los RNAt isodecodificadores en <i>Trityps</i>	62
Figura 26	Correlación entre el número de genes de RNAt y el uso de codones en <i>Trityps</i>	65
Figura 27	Estructuras secundarias de las secuencias de RNAr 5S en tripanosomátidos	67
Figura 28	Secuencias consenso de las cajas A, C y el Elemento Intermedio de los genes de RNAr 5S en <i>Trityps</i>	68

RESUMEN

Los tripanosomátidos son un grupo de protozoarios flagelados, algunos de los cuales causan enfermedades humanas que se transmiten por diferentes insectos vectores. En este grupo se incluye *Leishmania major*, agente causal de la leishmaniasis, *Trypanosoma brucei*, causante de la enfermedad de sueño africana y *Trypanosoma cruzi*, agente etiológico del mal de Chagas (conocidos como Trityps). Nuestro grupo de investigación está interesado en el estudio de la transcripción por la RNA polimerasa III (Pol III) en tripanosomátidos, que produce RNAs pequeños como los RNA de transferencia (RNAt) y el RNA ribosomal (RNAr) 5S. Los RNAt son moléculas cuya función consiste en descifrar los codones del RNA mensajero (RNAm) para acarrear el aminoácido correcto al ribosoma durante la síntesis de proteínas. El RNAr 5S es el más pequeño de los RNA que constituyen los ribosomas y, aunque su función no ha sido completamente entendida, podrían funcionar como un “puente” de comunicación entre diferentes sitios del ribosoma.

Para analizar las secuencias y la organización genómica de los genes de RNAt y de RNAr 5S se realizó un análisis *in silico* de los genomas de los Trityps. Primero, se usó el programa tRNA-Scan para confirmar la identidad de todos los genes de RNAt anotados en las bases de datos. Tras considerar algunas incongruencias encontradas en los genomas anotados de *T. brucei* y *T. cruzi* y después de incluir los genes de RNAt de selenocisteína (RNAt-Sec) (que no estaban anotados), nuestro análisis indicó la presencia de 83, 66 y 120 genes de RNAt en *L. major*, *T. brucei* y *T. cruzi*, respectivamente. La mayor parte de los genes están organizados en *clusters* de 2-10 genes en *L. major* y *T. brucei*. La mayoría de los genes en dichos *clusters* no muestran sintenia entre las tres especies de tripanosomátidos analizadas, es decir, no muestran un orden conservado de los genes. Sólo se encontraron cuatro *clusters* sinténicos, siendo el ejemplo más representativo un *cluster* de 13 genes de Pol III que se encuentra en el cromosoma 23 de *L. major*, el *cluster* 8-II en *T. brucei* y el *cluster* Tc6288 en *T. cruzi*.

Nuestro análisis reveló que los tripanosomátidos utilizan una estrategia de codificación conocida como “recambio de A1 - G1” para leer todos los codones en las proteínas, pues 16 anticodones no se encontraron en los RNAt de estos parásitos. Al igual que en otros eucariontes, en Trityps existe una correlación positiva entre el uso de codones y el número de isoaceptores de RNAt, lo cual fue evaluado utilizando la prueba de Spearman. Esto nos indicó que, de manera similar a otros eucariontes, en tripanosomátidos el patrón de uso de codones está adaptado a la abundancia de los RNAt. Por otro lado, se identificaron las secuencias consenso de los dominios de control interno (cajas A y B), y se analizaron las secuencias de término de la transcripción (*clusters* de 4 ó más Ts consecutivas) de los genes de RNAt. A este respecto, encontramos que *L. major* y

T. brucei presentan en promedio 5 ± 1 Ts como señal de término, mientras que *T. cruzi* presenta 7 ± 2 Ts; esta diferencia entre especies resultó estadísticamente significativa.

También se identificaron varios RNAt isodecodificadores en los Tritryps. Estos RNAt, recientemente descritos, poseen el mismo anticodón para el mismo aminoácido, pero presentan diferencias en el resto de la secuencia del gen. De manera interesante, algunas de las diferencias entre los genes isodecodificadores en Tritryps ocurren dentro, o muy cerca, de las secuencias promotoras internas, lo que sugiere que el control diferencial de la expresión de algunos RNAt isoaceptores en estos parásitos es factible. Asimismo, se presentan los resultados de los genes de RNAt con intrones y los RNAt-Met de iniciación y de elongación.

En cuanto al RNAr 5S, se encontraron 11, 8 y 120 genes en *L. major*, *T. brucei* y *T. cruzi*, respectivamente, mucho menos de lo encontrado para otros eucariontes. Mientras en *L. major* los genes se encuentran dispersos en el genoma (asociados a otros genes transcritos por Pol III), en *T. brucei* y *T. cruzi* se encuentran organizados en *clusters*. También se identificaron las estructuras secundarias del producto génico, así como los probables elementos promotores (cajas A y C y Elemento intermedio), resultados que concuerdan con los modelos propuestos en otros eucariontes.

ABSTRACT

The trypanosomatids are flagellated protozoan that cause devastating human diseases, which are transmitted by different insect vectors. *Leishmania major* produces leishmaniasis, *Trypanosoma brucei* is responsible of sleeping sickness in Africa, and *Trypanosoma cruzi* is the etiologic agent of Chagas disease; these three species are referred together as the Trityps. In addition to their medical importance, these parasites show unique genomic features. We are interested in the study of RNA polymerase III (Pol III) transcription in trypanosomatids. Pol III produces small essential RNAs such as transfer RNA (tRNA) and 5S ribosomal RNA (rRNA). The function of tRNA is to decipher the codons beared by the messenger RNA (mRNA) in order to carry up amino acids to the ribosome during protein synthesis. The 5S rRNA is the smallest RNA that is part of the ribosome and it apparently works as a “bridge” for communication among different sites of the ribosome, although its exact function has not yet been defined.

We performed an *in silico* study in order to analyze the sequences and genomic organization of tRNA and 5S rRNA genes in the Trityps. First, the tRNA-Scan program was used to confirm the identity of all tRNAs annotated in the databases. We found some misannotated tRNAs in the genomes of *T. brucei* and *T. cruzi*, as well as several previously unannotated selenocysteine (Sec) tRNA genes. Thus, our analysis indicated the presence of 83, 66 and 120 tRNA genes in *L. major*, *T. brucei* y *T. cruzi*, respectively. In *L. major* and *T. brucei*, most of the genes are organized in clusters of 2-10 genes with short intergenic sequences. While the majority of the tRNA clusters do not show synteny (conservation of gene order) between the Trityps, a cluster of 13 Pol III genes that is highly syntenic was identified. It is located on chromosome 23 from *L. major*, cluster 8-II in *T. brucei* and cluster Tc6288 in *T. cruzi*.

Also, the analysis of the isoacceptor tRNAs revealed that Trityps use the “A1 or G1 sparing strategy” as a decoding mode, in order to read all codons encoded by their proteins, since 16 anticodons were not found on the genomes of these parasites. Similarly to other eukaryotes, Trityps showed a positive correlation (Spearman coefficient) between the codon usage and the number of isoacceptor tRNAs, indicating that the codon usage pattern is coadaptated to the tRNA abundance in these parasites. Moreover, the internal control elements (boxes A and B) were analyzed in order to obtain consensus sequences for the putative promoters. The transcription termination signals (a cluster of at least 4 consecutive Ts) of the tRNA genes were also analyzed, showing an average of 5 Ts (± 1) in *L. major* and *T. brucei*, and 7 Ts (± 2) in *T. cruzi*.

We also report that the *Trypanosoma* species possess isodecoder tRNA genes, *i.e.* tRNAs having the same anticodon to recognize the same amino acid, but presenting differences elsewhere in the tRNA body. Interestingly, some of the differences between isodecoder genes occur in, or very close to, the promoter sequences, which suggests that differential control of expression of some isoacceptor tRNA genes in *Trypanosoma* is feasible. Furthermore, we present data related to tRNAs that hold an intron, as well as results concerning the elongator and initiator tRNA-Met genes.

Regarding the 5S rRNA, we found 11, 9 y 135 genes in *L. major*, *T. brucei* and *T. cruzi*, respectively; these numbers are smaller than those reported in other eukaryotic species. In *L. major* these genes are spread out in the genome, associated to other Pol-III transcribed genes. In *T. brucei* and *T. cruzi* the 5S rRNA genes are organized in a single cluster, in a specific location. The secondary structures of the 5S rRNAs were also determined, as well as consensus sequences for the putative promoter elements (boxes A and C and the Intermediate Element).

1. INTRODUCCIÓN

1.1. Generalidades sobre tripanosomátidos

Leishmania major, *Trypanosoma brucei* y *Trypanosoma cruzi* (conocidos como los TriTryps) son parásitos tripanosomátidos que causan graves enfermedades humanas. Estos parásitos, que pertenecen al grupo de los cinetoplástidos, poseen una organización genómica y estructuras celulares similares, y todos sufren cambios morfológicos durante sus ciclos de vida y son transmitidos a sus hospederos mamíferos por diferentes insectos vectores. Las enfermedades en humanos causadas por estos parásitos incluyen la enfermedad del sueño en África, que es causada por dos de las tres subespecies de *T. brucei*; la enfermedad de Chagas, causada por la infección con *T. cruzi*; y varias formas de leishmaniasis, causadas por diferentes especies de parásitos del género *Leishmania*.

Se calcula que 500 millones de personas, principalmente en regiones tropicales y subtropicales, están en riesgo de contraer alguna de estas enfermedades, y se ha estimado que más de 20 millones de individuos están infectados con alguno de estos patógenos, lo que resulta en más de 100 000 muertes por año. A la fecha, no se cuenta con un tratamiento efectivo para la cura de estas enfermedades. Además de la importancia médica, estos parásitos muestran numerosas particularidades metabólicas y estructurales (Stuart *et al*, 2008).

1.1.1. *Leishmaniasis*: importancia médica, epidemiología, manifestaciones clínicas.

La leishmaniasis es causada por ~20 especies patógenas de protozoarios del género *Leishmania*. El parásito es transmitido al hombre por un insecto vector, la mosca de la arena, del género *Phlebotomus* en el viejo mundo o *Lutzomya* en el continente Americano. Existen ~30 especies de insectos vectores que pueden diseminar la enfermedad. La leishmaniasis presenta diferentes manifestaciones clínicas: cutánea, mucocutánea y visceral, según la especie de *Leishmania* infectante (Tabla 1). Las dos primeras causan una discapacidad seria, dejando al individuo con cicatrices y lesiones permanentes que afectan su desenvolvimiento social, mientras que la última puede ser mortal (Murray *et al*, 2005).

Leishmaniasis cutánea: esta forma de la enfermedad produce úlceras en la piel de la cara, brazos y piernas que pueden llegar a sanar espontáneamente (Fig. 1C). Estas lesiones producen

cicatrices desfigurantes que afectan la estética del individuo. Una variante es la leishmaniasis cutánea difusa, cuyas lesiones son muy diseminadas y devastadoras para el paciente (Tabla 1).

Leishmaniasis mucocutánea: las lesiones conllevan a la destrucción total o parcial de las membranas mucosas de la nariz, boca y faringe así como de los tejidos vecinos con mutilación de la cara (Fig. 1D y Tabla 1).

Leishmaniasis visceral: también conocida como *kala azar* (enfermedad negra) en la India, se caracteriza por brotes irregulares de fiebre, gran pérdida de peso y anemia. La principal manifestación es la gran inflamación de vísceras como el bazo e hígado, lo cual abulta considerablemente el abdomen (Fig. 1E, Tabla 1). Otros órganos afectados son la médula ósea y los ganglios linfáticos. La piel está hiperpigmentada, signo que originó el nombre de la enfermedad en la India. Esta forma de la enfermedad presenta un grado muy alto de mortalidad (Murray, 2005, Desjeux, 2004).

Tabla 1. Especies de *Leishmania* y la forma de leishmaniasis que producen

Forma de la enfermedad	<i>Leishmania spp</i>	Distribución
Cutánea	<i>L. mexicana</i> *, <i>L. amazonensis</i> , <i>L. garnham</i> , <i>L. peruviana</i> , <i>L. braziliensis</i>	Nuevo Mundo
	<i>L. major</i> *, <i>L. tropica</i> , <i>L. aethiopica</i>	Viejo Mundo
	<i>L. infantum</i> , <i>L. chagasi</i>	Mediterráneo y región del Mar Caspio
Mucocutánea	<i>L. braziliensis</i> , <i>L. panamensis</i> , <i>L. guyanensis</i>	Nuevo Mundo
	<i>L. donovani</i> , <i>L. major</i> (<i>L. infantum</i> en pacientes inmunosuprimidos)	Viejo Mundo
Visceral	<i>L. chagasi</i>	América (Brasil, Venezuela, México, Colombia, Bolivia, Argentina, Centroamérica)
	<i>L. donovani</i>	Asia (India, China)
	<i>L. infantum</i>	Mediterráneo, norte de África

(* Lesiones que sanan espontáneamente)

Actualmente existen 350 millones de personas en 88 países alrededor del mundo en riesgo de contraer esta enfermedad, 12 millones de personas infectas y alrededor de 70 000 muertes por año. La incidencia anual está estimada en 1- 1.5 millones de casos de leishmaniasis cutánea y 500 000 de visceral (Desjeux, 2004, Murray, 2005).

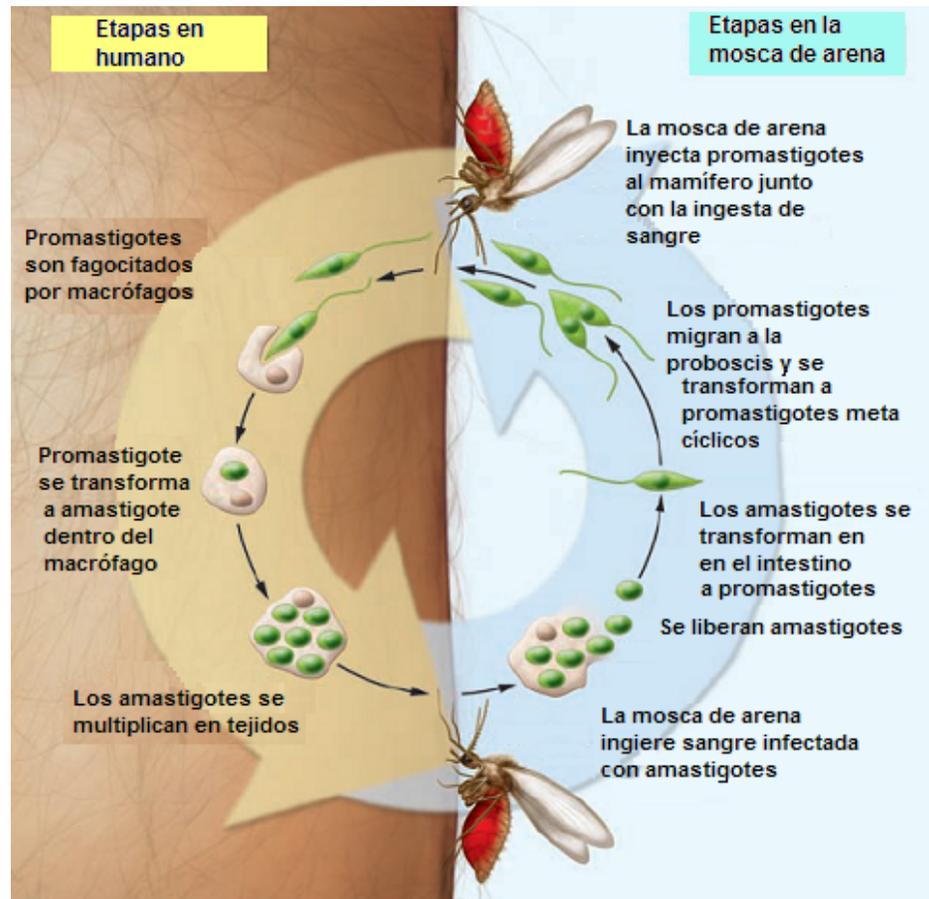
1.1.1.1. Ciclo de vida de *Leishmania*

Los parásitos del género *Leishmania* spp son transmitidos por moscas de la arena hembras de los géneros *Lutzomya* y *Phlebotomus*, previamente infectadas con estos protozoarios (fig. 1A). Durante la ingesta de sangre, las moscas inoculan en el mamífero (humano u otros reservorios) promastigotes metacíclicos (fig. 1B), la forma infecciosa del parásito. Los promastigotes metacíclicos son fagocitados por los macrófagos y otras células fagocíticas y dentro de los fagolisosomas (o vacuolas parasitóforas) se diferencian a amastigotes (fig. 1B), la forma pequeña y redondeada del parásito, responsable de la invasión a tejidos del mamífero, que se replica por fisión binaria. Nuevas moscas de la arena toman macrófagos infectados durante la ingesta de sangre. En el intestino de las moscas los amastigotes se transforman a promastigotes procíclicos (fig. 1B), una forma no infectiva y altamente replicativa del parásito. Estas formas se unen a la pared del intestino medio para luego liberarse y migrar al aparato bucal de la mosca, diferenciándose a promastigotes metacíclicos no replicativos que son transmitidos por el insecto al mamífero durante otra ingesta de sangre (Fig. 1A) (Desjeux, 2004, Stuart *et al*, 2008).

1.1.2. Aspectos generales de *Trypanosoma brucei*

T. brucei es el agente causal de la enfermedad del sueño en África. La primera etapa de la enfermedad (etapa hemolinfática) está acompañada de fiebre, dolor de cabeza, adenopatía y prurito. Las variaciones antigénicas del parásito permiten la evasión a la respuesta inmune del hospedero, lo que produce picos de parasitemia. La etapa crónica de la enfermedad (etapa del sistema nervioso central) se acompaña de síntomas neurológicos severos que incluyen anomalías mentales, sensoriales y del sueño. La tripanosomiasis africana se presenta en la región subsahariana en territorios de la mosca tsetse. *T. brucei gambiense* causa primordialmente la forma crónica de la enfermedad en las regiones centrales y del oeste del África subsahariana, mientras que *T. brucei rhodesiense* causa principalmente la forma aguda en regiones del este y sur de la zona. *T. brucei brucei* por su parte, es patógeno para el ganado, no así para el humano. Mas de 20 especies de moscas tsetse (*Glossina spp*) pueden transmitir estos parásitos. Existen en la actualidad ~50 millones de personas en riesgo de contraer la enfermedad, 80 000 personas infectadas, presentándose ~ 30 000 muertes por año (Stuart *et al*, 2008)

A



B

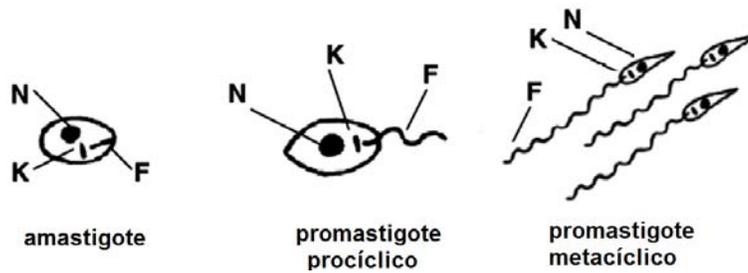


Figura 1. Ciclo de vida de *Leishmania* spp y manifestaciones clínicas que produce (pie de figura siguiente página)

Figura 1. Ciclo de vida de *Leishmania spp* y manifestaciones clínicas que produce. A). Ciclo de vida de *Leishmania spp* (tomado y modificado de Stuart *et al*, 2008). Ver texto para detalles. B) Etapas de desarrollo de *Leishmania spp*. Se muestra la morfología del amastigote, promastigote procíclico y promastigote metacíclico. Cada forma posee un núcleo (N), cinetoplasto (K) y flagelo (F). El flagelo en amastigotes es interno y no funcional, mientras que en los promastigotes el flagelo se extiende a través de la célula y emerge por la parte anterior (tomado de Bates, 2007), 1C) Lesiones por leishmaniasis cutánea, 1D) leishmaniasis mucocutánea desfigurante, 1E) leishmaniasis visceral (tomado de Murray *et al*, 2005).

El ciclo de vida inicia cuando la mosca ingiere tripanosomas sanguíneos junto con la ingesta de sangre contaminada de mamíferos infectados. Los parásitos se transforman a tripomastigotes procíclicos en el intestino del vector y luego a epimastigotes en las glándulas salivales para finalmente convertirse en tripomastigotes metacíclicos, forma infectiva que se transmite a otro mamífero en una nueva ingestión de sangre. Esta forma se diferencia a tripomastigotes sanguíneos, que se multiplican en varios fluidos corporales como sangre y líquido espinal. En la etapa crónica, los parásitos invaden el sistema nervioso central (Stuart *et al*, 2008).

1.1.3. Aspectos generales de *Trypanosoma cruzi*

T. cruzi es el agente causal de la tripanosomiasis americana o enfermedad de Chagas, que se presenta a lo largo del continente Americano. El parásito es transmitido por chinches reduvidas de los géneros *Rhodnius*, *Triatoma* y *Pastronylus*. La enfermedad se caracteriza por tres fases: a) aguda, con reacciones inflamatorias en el sitio de infección, fiebres, anemia, miocarditis aguda, etc., b) indeterminada, sin signos de enfermedad, aparentemente asintomática y c) crónica, con manifestaciones de enfermedad cardíaca o visceromegalias (hígado, bazo, estómago, intestinos). Actualmente, existen alrededor de 100 millones de personas en América Latina en riesgo de contraer esta enfermedad, 8-11 millones de personas infectadas y ~ 14 000 muertes por año (Stuart *et al*, 2008).

T. cruzi tiene un ciclo de vida bifásico, en el que los vectores se infectan al ingerir sangre de mamíferos infectada con tripomastigotes; éstos se diferencian en el tubo digestivo del vector a epimastigotes y posteriormente a tripomastigotes metacíclicos, siendo éstos últimos los infectantes para el hospedero vertebrado. Al picar nuevamente al hombre o a algún otro vertebrado, los triatomíneos infectados defecan, depositando tripomastigotes metacíclicos junto con las heces sobre la superficie de la piel. Los tripomastigotes pueden introducirse al organismo a través de la picadura o heridas de la piel o atravesando la mucosa ocular, nasal o bucal. En el mamífero, los tripomastigotes metacíclicos son fagocitados por macrófagos, en los que se transforman en amastigotes, la forma intracelular. Los amastigotes se multiplican y diferencian a tripomastigotes,

que se liberan al torrente sanguíneo e infectan múltiples órganos. Los tripomastigotes pueden ser tomados por un nuevo insecto durante la ingesta de sangre (Burleigh y Woolsey, 2002).

1.2. Particularidades biológicas y organización genómica de tripanosomátidos

Los tripanosomátidos poseen numerosas particularidades metabólicas y estructurales, por ejemplo: 1) los glicosomas, estructuras citoplasmáticas que contienen enzimas glicolíticas e intermediarios de la glicólisis, 2) el cinetoplasto, DNA mitocondrial conformado por círculos concatenados, el cual equivale al 10% del DNA celular total y que contiene maxicírculos (codificantes de genes mitocondriales) y minicírculos (conteniendo los RNA guías) 3) la edición del RNAm mitocondrial, es decir, la modificación post-transcripcional de los RNA mensajeros mitocondriales por la adición y la eliminación de uridinas, 4) la expresión policistónica de los genes nucleares, 5) el procesamiento y maduración de los RNAm por *trans-splicing* (adición del miniexón de 39 pb al extremo 5' y 6) la variación antigénica de las proteínas variantes de superficie (VSGs) de *T. brucei* para la evasión de la respuesta inmune del hospedero (Campbell *et al*, 2003).

La publicación de los genomas completos de *L. major*, *T. brucei* y *T. cruzi* (Ivens, *et al*, 2005; Berriman, *et al*, 2005; El-Sayed, *et al*, 2005, respectivamente) muestra que los genes en todos los cromosomas de estos parásitos están organizados en *clusters* largos de genes direccionales (*Directional gene cluster*, DGCs), que son grupos de genes codificadores de proteínas sin una función común localizados consecutivamente en la misma cadena de DNA (Hernández-Rivas *et al*, 2007). El genoma nuclear de *L. major* (32.8 megabases, Mb) se reparte entre 36 cromosomas que varían de 0.28 a 2.8 Mb. *T. cruzi* posee ~28 cromosomas y *T. brucei* 11 megacromosomas entre 1 y 6 Mb (más numerosos minicromosomas con secuencias repetitivas). El número de genes codificadores de proteínas por genoma haploide se ha reportado de ~8 200, ~12 000 y ~8 100 en *L. major*, *T. brucei* y *T. cruzi*, respectivamente, aunque el número exacto de genes no se ha determinado para *T. cruzi*. Estos y otros datos se presentan en la tabla 2. Los tripanosomátidos son diploides, aunque algunos cromosomas son aneuploides.

A pesar de haberse separado en la evolución hace más de 200 millones de años, se puede observar una fuerte conservación del orden de los genes (sintenia) en el genoma de los tripanosomátidos. Las regiones altamente sinténicas comprenden el 94% de los 6 200 genes conservados entre las tres especies y representan 68 y 75% del genoma de *T. brucei* y *L. major*, respectivamente (Hernández-Rivas *et al*, 2007, El-Sayed *et al*, 2005).

Tabla 2. Resumen de los genomas de los tripanosomátidos.

	<i>L. major</i>	<i>T. brucei</i>	<i>T. cruzi</i>
Tamaño del genoma (Mb)	32.8	26	55
G+C % (genoma)	59.7	46.4	51
Cromosomas	36	11 §	28 ♦
GCP*	8 233	8164	~12 000
RNA transferencia	83	65	115
RNA ribosomal	63 ¥	56	219

* Genes Codificadores de Proteínas por genoma haploide, sin incluir pseudogenes

¥ El número exacto no se ha determinado, § considerando sólo los megacromosomas, ♦ el número exacto no se conoce y los homólogos pueden variar considerablemente en tamaño (tomado parcialmente de El-Sayed *et al*, 2005)

1.3. Generalidades sobre la transcripción por RNA polimerasa III (Pol III)

La transcripción de genes en eucariontes está a cargo de tres RNA polimerasas (Pol): Pol I que sintetiza RNA ribosoma (RNAr) 18S, 28S y 5.8S; Pol II que sintetiza los RNA mensajeros (RNAm) codificantes de proteínas y RNAs no codificadores (RNAnc) como los pequeños nucleares (RNAsn), pequeños nucleolares (RNAsno) y microRNAs (miRNA). Pol III es la mayor de las RNA polimerasas, pues pesa alrededor de 700 kDa y posee 17 subunidades. Todos sus productos son transcritos cortos como el RNAr 5S, los RNA de transferencia (RNAt), el RNA 7SL (componente de la partícula de reconocimiento del péptido señal) y el RNAsn U6. Además, recientemente se ha identificado un microRNA de humano transcrito por Pol III. En eucariontes, los RNA codificados por Pol III se encuentran generalmente en familias multigénicas, en donde cada gen es transcrito independientemente. Otros productos de Pol III están listados en la tabla 3 (Dieci *et al*, 2007):

Tabla 3. Algunos RNAs sintetizados por Pol III.

RNA	Descripción
MRP	Procesamiento de pre-RNAr
H1	Procesamiento de RNAt
<i>Vault</i> RNA (VA RNA)	Presentes en RNPs citoplasmáticos implicados en resistencia de tumores a fármacos, ensamblaje y/o transporte macromolecular
7SK	Identificado sólo en vertebrados hasta la fecha, son represores de la elongación por Pol II al unirse al factor P-TEFb
BC1 y BC200	Identificados en roedor y primates, respectivamente, parecen tener un papel importante en la traducción de RNAm de dendritas.

Tabla 3. Continuación

SINEs (<i>Short interspersed repeated DNA elements-encoded RNAs</i>)	Retrotransposones no autónomos originados por genes transcritos por Pol III. Parecen regular la expresión génica a nivel post transcripcional, modular la traducción de proteínas o apagar su expresión (RNAs Alu). Existen más de un millón de SINEs en el genoma humano, la mayoría inactivos.
sRNA76	RNA <i>t-like</i> identificado en tripanosomátidos y que coprecipita con el RNA 7SL en la partícula de reconocimiento de señal SRP (Liu, L., <i>et al.</i> 2003)

1.3.1. Elementos Promotores de la Transcripción por Pol III

La característica más distintiva e inusual de los promotores de Pol III es que la mayoría de ellos requiere secuencias localizadas “río abajo” del sitio de inicio de la transcripción (SIT, +1), dentro de la región transcrita. Existen tres tipos diferentes de promotores para Pol III (fig. 2):

a) Tipo 1: requieren tres elementos internos para una transcripción eficiente: una caja A, una caja C y un elemento intermedio (EI). Este tipo de promotor se encuentra en los genes de RNAr 5S. Se ha reportado que en *Saccharomyces cerevisiae* sólo se requiere la caja C para la transcripción. La caja A está localizada entre +50 y +64, el EI entre +67 a +72 y la caja C de +80 a +97.

b) Tipo 2: se han encontrado en los genes RNAt, y se componen de las cajas A (cerca del SIT, entre +11 y +21) y B (de localización variable), las cuales son secuencias altamente conservadas que se localizan en el asa D y T, respectivamente (Figura 5), que son importantes para la función del RNAt. Es importante mencionar que se ha encontrado un elemento tipo caja TATA localizado río arriba de algunos genes de RNAt, la cual es esencial para la expresión de estos genes.

c) Tipo 3: formados por elementos que residen exclusivamente “río arriba” del gen; éstos son una caja TATA (entre -30 y -25), una secuencia proximal (PSE, *proximal sequence element*, entre -66 y -47), que constituyen el promotor basal, y una secuencia distal (DSE, *distal sequence element*, entre -244 y -214), que potencializa la transcripción. Ejemplos de este tipo de promotor se encuentran en los genes del RNAsn U6 humano (componente del spliceosoma), el RNA 7SK, y el H1 (Schramm y Hernandez, 2002; Paule y White, 2000).

Algunos genes transcritos por Pol III no caen dentro de ninguno de los tipos 1, 2 ó 3, ya que usan combinaciones de elementos promotores. Aunque la caja TATA es un elemento extragénico característico de los promotores tipo 3, también ha sido encontrado en las regiones 5' de algunos genes con promotores internos, por ejemplo, algunos genes de RNAt en plantas, *Drosophila* y *Schizosaccharomyces pombe*. Un caso interesante es el de *S. pombe*, organismo en el que prácticamente todos sus RNAt y RNAr 5S poseen una caja TATA río arriba del SIT (Schramm y Hernandez, 2002). Otros ejemplos de genes que requieren elementos tanto internos como externos para su transcripción y no caen plenamente en alguno de los tipos de promotor incluyen los genes de RNAt de selenocisteína (RNAt-Sec, ver sección 1.6), el 7SL de humano y el RNA *vault* de rata (Paul y White, 2000).

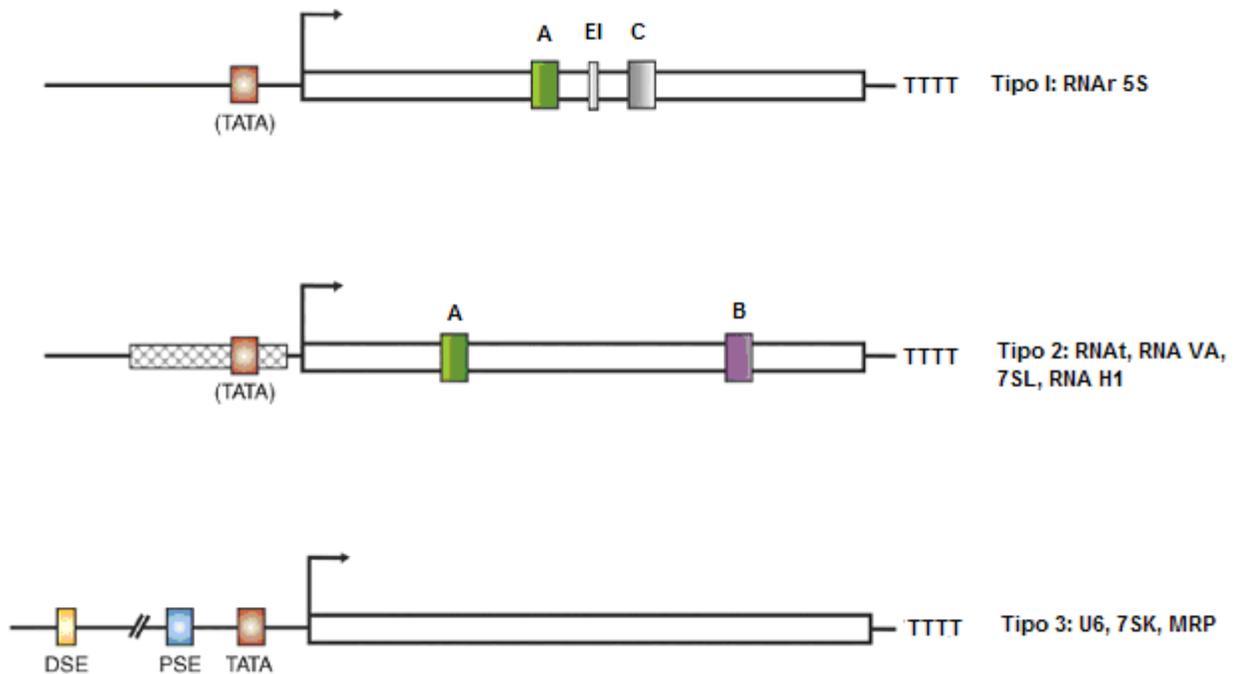


Figura 2: Los tres tipos de promotores de Pol III. Se indican las posiciones aproximadas de los elementos de control externos e internos. El tipo 1 posee las cajas A, C y el elemento intermedio (EI). El tipo 2 requiere de una caja A y una caja B. El promotor tipo 3 posee elementos extragénicos que incluyen un elemento proximal (PSE) y una caja TATA como elementos promotores basales y un elemento distal (DSE), el cual optimiza el nivel de transcripción. En algunos organismos se han observado cajas TATA asociadas a promotores tipo 1 y 2 (Dieci, G., *et al*, 2007).

1.3.2. Factores de transcripción asociados a la RNA polimerasa III

El factor TFIIIB es un factor de transcripción común para todos los promotores de Pol III. En *S. cerevisiae*, TFIIIB es un complejo de tres polipéptidos: TBP (TATA *binding protein*), B'' (Bdp1 en humanos) y BRF (*TFIIB-related factor*). En los promotores tipo 2, el factor TFIIIC reconoce las cajas A y B; esta unión proteínas-DNA permite posteriormente el reclutamiento del factor TFIIIB. Se ha mostrado que varias subunidades de TFIIIC se extienden a lo largo del gen de RNAt-Tyr. Al parecer es la caja B el determinante principal para la unión de dicho factor, aunque la caja A es también contactada (Paul y White, 2000). A su vez, la unión de TFIIIB al promotor favorece el reclutamiento de Pol III (Fig. 3A) (Schramm y Hernandez, 2002). En los promotores de tipo 1, las cajas A, C y el elemento intermedio son reconocidos por el factor TFIIIA, proteína con nueve dedos de zinc. La caja C es reconocida por tres dedos del extremo N-terminal, mientras que los dedos 7-9 contactan la caja A. El complejo TFIIIA-DNA dirige a TFIIIC hacia el promotor de RNAr 5S. Tras la unión de TFIIIC, la vía de reclutamiento de la polimerasa es similar a la que ocurre con los promotores tipo 2, con el reclutamiento de TFIIIB y Pol III (Fig. 3B) (Schramm y Hernandez, 2002; Paul y White, 2000).

En los promotores tipo 3, el PSE es reconocido por el complejo multiproteico PTF (factor de transcripción PSE, *PSE transcription factor*), también llamado SNAPc (*snRNA activating protein complex*, complejo activador de RNAsn). Por su parte, la caja TATA es reconocida por una variante de TFIIIB, en la que Brf1 es reemplazado por Brf2. El DSE de los promotores tipo 3 posee un octámero en su secuencia, sitio de unión de Oct-1 y una Secuencia SPH (*Sph1 postoctamer homology*) que recluta Staf/ZNF/SBF. Las interacciones DNA-proteína son reforzadas por interacciones proteína-proteína entre SNAPc y TBP. La unión de SNAPc y el factor TFIIIB-like conducen al reclutamiento de Pol III, probablemente a través de contactos proteína-proteína con los factores mencionados (Fig. 3C) (Schramm y Hernandez, 2002, Paul y White, 2000).

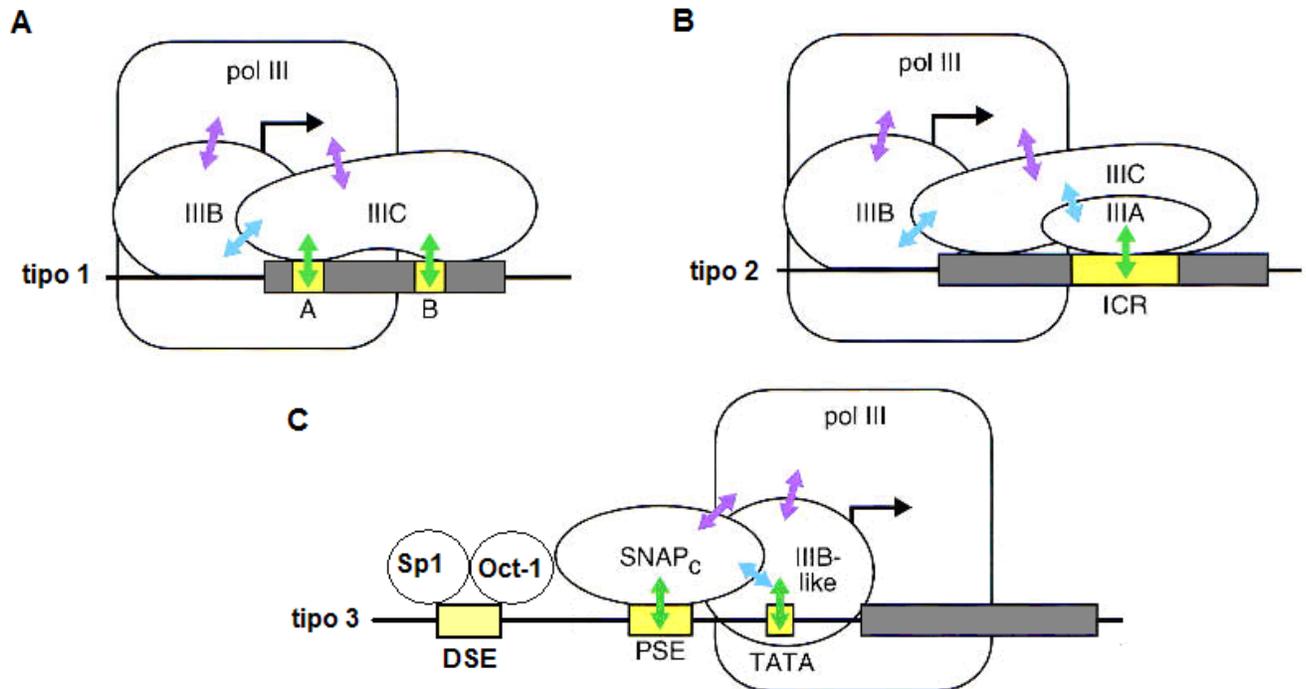


Figura 3: Diferentes vías para el reclutamiento del TFIIIB y RNA polimerasa III. Se muestran los complejos de iniciación formados en los promotores tipo 2 (A), tipo 1 (B) y tipo 3 (C). Las barras grises representan los genes transcritos por Pol III, las regiones amarillas los promotores correspondientes. Las flechas verdes simbolizan las interacciones de DNA-proteínas con los elementos promotores, las flechas azules contactos proteína-proteína entre factores de transcripción, y las flechas púrpuras señalan contactos proteína-proteína entre RNA pol III y factores de transcripción (tomado de Schramm y Hernandez, 2002).

1.3.3. Terminación de la transcripción por Pol III

En eucariontes, una secuencia de cuatro o más residuos de T localizada al extremo 3' de los genes actúa como señal para la terminación de la transcripción por Pol III (Fig. 2). En células humanas ya se han reportado varios factores involucrados en la terminación de la transcripción: el Factor Positivo 4 (PC4), Factor Nuclear 1 (NF1), el antígeno La y una proteína de unión al extremo UUU-OH. Algunas subunidades de TFIIIC, como TFIIIC2, parecen estar involucradas en el término y reinicio de la transcripción. Se ha reportado que las subunidades de Pol III C11, C37 y C53 forman un complejo que participa también en el término y reinicio de la transcripción (Hernández-Rivas *et al*, 2007). La fuerza de la secuencia de Ts como terminador está altamente influenciada por las secuencias río arriba y río abajo del tracto de Ts. Por ejemplo, en *S. cerevisiae*, la secuencia CT río abajo del tracto de Ts actúa como un elemento que debilita el efecto de terminador, mientras que la terminación de la transcripción es eficiente cuando las Ts están seguidas por A o G (Braglia, 2005).

1.4. Generalidades del RNA ribosomal 5S

El RNAr 5S es el RNA ribosomal más pequeño, con aproximadamente 120 nucleótidos y una masa de ~40 kDa. El RNAr 5S se encuentra en todos los ribosomas, excepto en los de las mitocondrias de ciertos hongos, de animales superiores y de la mayoría de los protistas (Barciszewska *et al*, 2000). Específicamente, el RNAr 5S se encuentra en la subunidad mayor eucarionte (60S), que además posee también los RNAr 28S y 5.8S, así como ~50 proteínas.

1.4.1. Función del RNA ribosomal 5S

La función exacta del RNAr 5S no se conoce completamente hasta el momento. En la subunidad mayor, este RNAr se localiza cerca de los sitios de unión a la peptidil transferasa y a los factores de traducción (Barciszewska *et al*, 2000). Estudios en levaduras y en *E. coli* han sugerido algunas probables funciones del RNAr 5S: i) el RNAr 5S es crucial para la unión del RNAt al sitio A del ribosoma (sitio de unión aminoacil-RNAt); ii) el RNAr 5S sirve como conexión física entre diferentes regiones del ribosoma, por lo que podría ser un transductor de señales para la comunicación entre dichas regiones, ayudando así a coordinar los eventos de traducción. El sitio P del ribosoma (peptidil-transferasa) y el sitio de unión del factor EF-G podrían estarse comunicando a través del 5S y otros RNAr de la subunidad mayor (Hellman *et al*, 2007); iii) el RNAr 5S podría también funcionar estabilizando a la subunidad mayor. Cabe mencionar que el 5S por si mismo no contacta directamente los RNAt unidos a ninguno de los sitios P o A; ni tampoco es un componente de los centros de posicionamiento de la peptidiltransferasa ni de los factores de elongación (Dinman, 2005).

1.4.2. Estructura secundaria del RNAr 5S

La estructura secundaria del RNAr 5S consiste de cinco hélices (I-V), dos asas tipo *hairpin* (C y D), dos asas internas (B y E) y una región “bisagra” (A); regiones organizadas en una estructura de tres hélices (Fig. 4) (Szymanski *et al*, 2003). La región del asa E del RNAr 5S es una región de doble hélice que parece ser importante para las interacciones específicas RNA-proteínas (Barciszewska *et al*, 2000). La hélice II y el asa E presentan protuberancias conformadas por un nucleótido. La hélice III tiene de dos a tres bases altamente conservadas en su protuberancia y, en *Xenopus laevis*, las As que se encuentran en esta región parecen ser importantes para la unión del 5S a la proteína ribosomal L5. La hélice IV posee pares de bases G-C y G-U (Fig. 4) (Szymanski *et*

al, 2003). Acerca de su estructura terciaria, la cristalización del RNAr 5S de *Thermus flavus* permitió observar que la hélice II forma un sitio de unión para proteínas ribosomales y el factor TFIIIA.

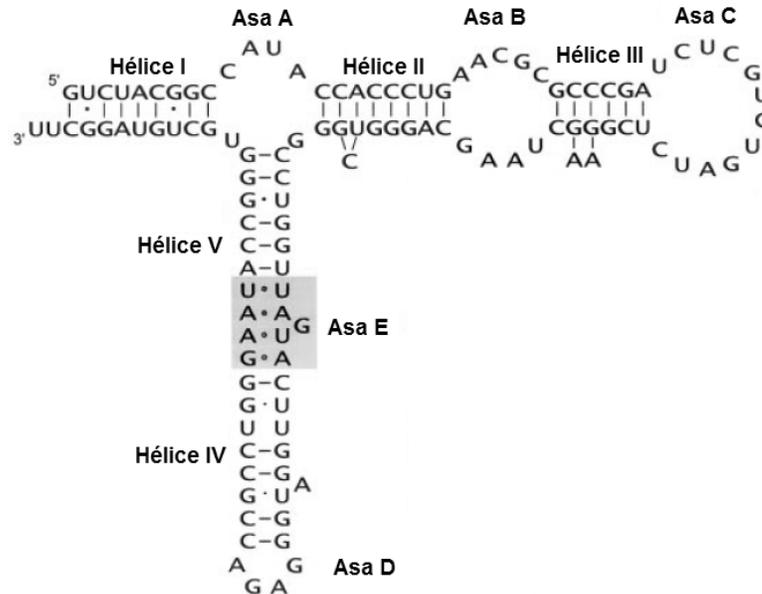


Figura 4. Estructura Secundaria del RNAr 5S humano. La estructura consiste en cinco hélices, regiones de doble cadena (I-V) y cinco asas (A-E). El asa E, que difiere entre eucariontes y bacterias está resaltada en sombreado (tomada de Szymanski *et al*, 2003).

1.4.3. Distribución en el genoma y expresión del RNAr 5S

En eucariontes superiores, el número de genes de RNAr 5S varía ampliamente encontrándose hasta varios miles en vertebrados y plantas (Barciszewska *et al*, 2000). Los genes de RNAr 5S generalmente no se encuentran contiguos al resto de los genes de RNA ribosomal, excepto en *S. cerevisiae* y *Dictyostelium discoideum*, en los que los genes 5S se encuentran adyacentes a los repetidos de DNAr. En la mayoría de las especies los genes 5S se encuentran en un *cluster* mayor de repetidos en tándem, como en *Gallus gallus*, o en pocos *clusters* pequeños como en *Arabidopsis thaliana* y *Drosophila virilis*. En pocos casos los genes 5S se encuentran esparcidos de manera lineal como en *S. pombe*, *Neurospora crassa* y *Acanthamoeba castellanii* (Haeusler y Engelke, 2006).

1.5. Generalidades sobre RNAs de transferencia

1.5.1. Estructura y función de los RNAs de transferencia

La función de los RNAt es descifrar los codones del RNAm. Cada tipo de aminoácido tiene su propio RNAt, que se une al aminoácido y lo acarrea al extremo de la cadena polipeptídica en crecimiento durante la síntesis de proteínas. Durante la traducción, el RNAt se une químicamente a su aminoácido correspondiente formando un aminoacil-RNAt. El RNAt correcto con su aminoácido unido es seleccionado en base a su anticodón, una secuencia de tres nucleótidos que se aparea con su codón complementario en el RNAm.

Las moléculas de RNAt tienen de 70 a 90 nt de longitud y su función depende de su estructura tridimensional. En la estructura secundaria, el RNAt se ensambla en un arreglo tallo-asa similar a una hoja de trébol (Fig. 5A). Los cuatro tallos o brazos son dobles hélices cortas con apareamientos de bases Watson-Crick. Tres de los cuatro tallos (brazo D, brazo T ψ C, brazo del anticodón) tienen asas de siete u ocho nucleótidos en sus extremos, mientras que el cuarto tallo carece de asa y contiene los extremos 3' y 5' de la cadena (brazo aceptor). Visto en tres dimensiones, la molécula de RNAt se pliega en forma de L con el asa del anticodón y el tallo aceptor formando los extremos de los dos brazos (Fig. 5B).

El brazo aceptor está formado por aproximadamente 7 pares de bases (pb) y posee los extremos 5' y 3' de la molécula. El aminoácido se encuentra unido al extremo 3' del RNAt, a la adenosina de la secuencia terminal CCA (Fig. 5A). El brazo D recibe su nombre por el nucleótido modificado dihidouridina, casi siempre presente en el asa D. La región del brazo y asa D en el DNAt posee la secuencia correspondiente a la caja A, elemento promotor reconocido por el factor de transcripción TFIIC (ver Sección Transcripción por Pol III).

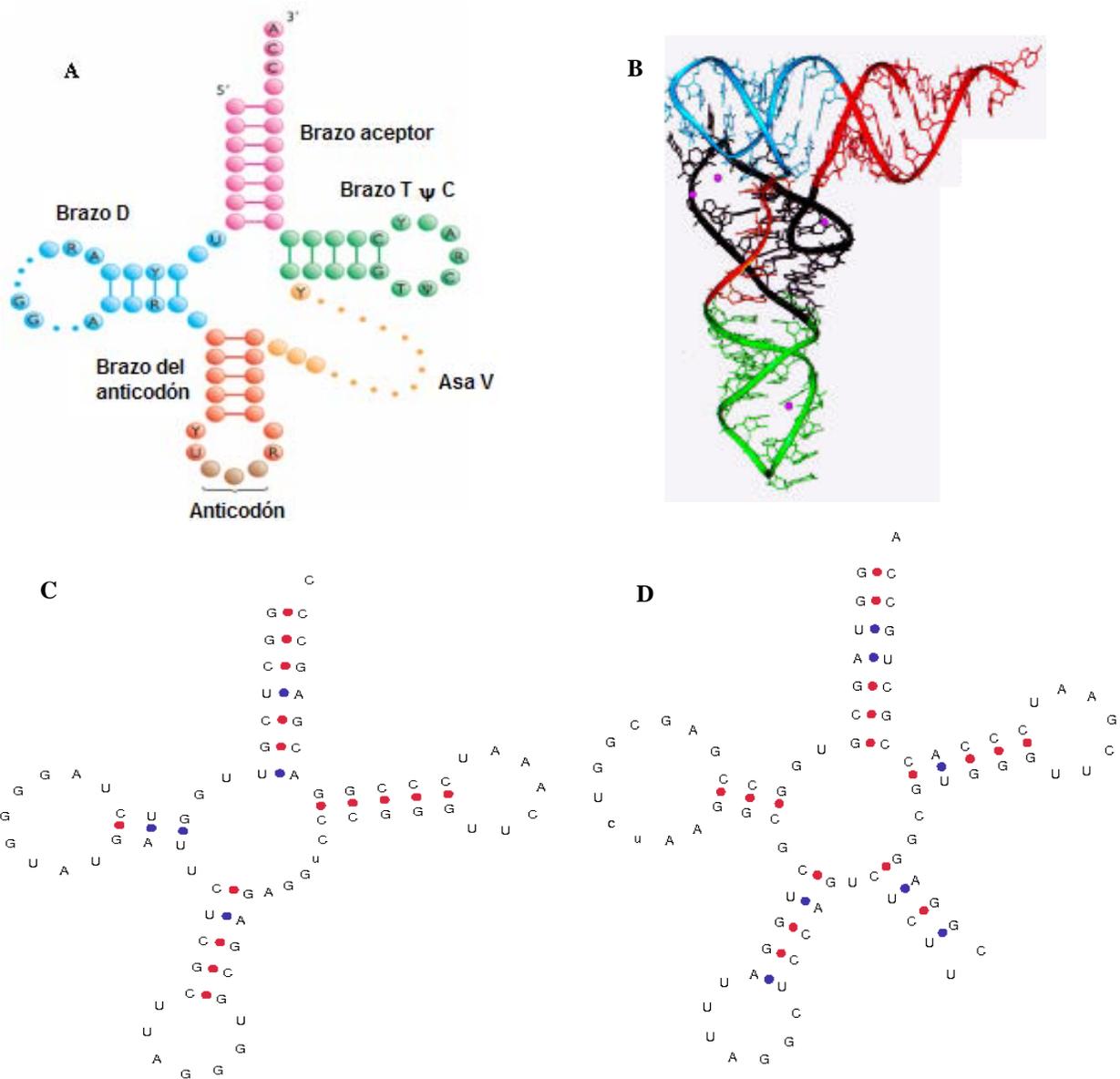


Figura 5. Estructura del RNAt: A) Estructura Secundaria. En rosa el brazo aceptor que se une al aminoácido correspondiente, en azul, el brazo y asa D, en naranja el brazo y asa del anticod3n y en verde, el brazo T ψ C. El brazo V, variable se presentan en amarillo y sólo forma parte de algunos RNAt, B) Estructura terciaria del RNAt, plegado en forma de L, C) Estructura Secundaria de un RNAt-Pro (clase I), D) Estructura Secundaria del RNAt-Leu (clase II), con el brazo V característico.

El anticod3n, que se encuentra en el brazo del mismo nombre, forma un apareamiento de bases con el RNAm durante la traducci3n. En el RNAt maduro, A en la primera posici3n del anticod3n es siempre modificada a inosina (I) por la enzima tRNA A34-desaminasa. El brazo T ψ C recibe su nombre por contener siempre ribotimidina (T) y pseudouridina (ψ) en el RNAt maduro. Esta regi3n en el DNAt contiene el segundo promotor interno de los genes: la caja B, secuencia de uni3n al factor de transcripci3n TFIIC (Fig. 5A) (Marck y Grosjean, 2002).

El brazo V o brazo variable se encuentra entre el brazo del anticodón y el brazo T ψ C (Fig. 5A). Dependiendo de la longitud del brazo variable, pueden distinguirse dos clases de RNAt: a) RNAt clase I: RNAt con un asa variable corta de 4-5 residuos, aunque pueden encontrarse 8 ocasionalmente (Fig. 5C) y b) RNAt clase II: aquellos RNAt que tienen un brazo variable largo, de 11-23 residuos (en la que el asa tiene de 3-5 residuos, Fig. 5D). Los tRNA clase II son únicamente los RNAt de leucina y serina; el resto son de clase I (Brennan y Sundaralingam, 1976, Marck y Grosjean, 2002). En el RNAt-Ser, el brazo variable es un determinante para su reconocimiento por las RNAt-seril sintetasas (Marck y Grosjean, 2002).

En algunos casos, los genes de RNAt poseen intrones, los cuales se localizan siempre en la misma posición, es decir, entre los nucleótidos 37 y 38 del asa del anticodón, un nucleótido río abajo del anticodón. La longitud de los intrones es muy variable, ya que va desde 7 a 121 bases en procariontes; entre eucariontes, el intrón más largo encontrado es de 60 pb. Se ha observado que los intrones en los DNAt son más frecuentes en eucariontes inferiores que en los superiores. No se conoce a la fecha la función exacta de los intrones en los DNAt, sin embargo, se ha demostrado que el intrón tiene un papel esencial en la maduración química y estructural del pre-RNAt (Marck y Grosjean, 2002).

1.5.2. RNAt isoaceptores

Existen RNAt que acarrean el mismo aminoácido, pero poseen diferentes anticodones; éstos se conocen como **RNAt isoaceptores**. El código genético universal posee 61 codones para codificar los 20 aminoácidos universales, así como tres codones de paro. Así, la mayoría de los aminoácidos están codificados por más de un codón, lo que se conoce como la “degeneración” del código genético. Nueve aminoácidos están codificados por dos codones, cinco aminoácidos poseen cuatro codones y tres aminoácidos tienen seis codones. Un aminoácido tiene tres codones y sólo dos aminoácidos tienen codones únicos (Fig. 6) (Agris *et al*, 2007).

		SEGUNDA								
		U		C		A		G		
PRIMERA	U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
		UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C
		UUA	Leu	UCA	Ser	UAA	Paro	UGA	Paro	A
		UUG	Leu	UCG	Ser	UAG	Paro	UGG	Trp	G
	C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
		CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C
		CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A
		CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G
	A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
		AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C
		AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A
		AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg	G
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U	
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C	
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A	
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G	

Figura 6. El código genético. Los códigos de aminoácidos que están dos veces degenerados se muestran en gris, los cuatro veces degenerados en amarillo y los seis veces degenerados en azul. En verde la familia de tres codones de Ile, y los codones únicos de Met y Trp en blanco. Los codones de paro en rojo (tomada de Agris, P. F. *et al*, 2007). Del lado izquierdo, la primera letra del codón, en la parte superior, la segunda y del lado derecha la tercera posición o posición *wobble*.

Los seres vivos pueden leer los 61 codones universales con mucho menos de 61 isoaceptores y se ha descrito que en los genomas eucariontes se encuentran únicamente de 41 a 55 RNAt isoaceptores (Goodenbour y Pan, 2006). Francis Crick propuso un mecanismo para explicar cómo los RNAt pueden leer más de un codón: la Hipótesis *Wobble*, que propone que las dos últimas bases del anticodón en el RNAt se aparean específicamente con las dos primeras del codón en el RNAm, es decir, son apareamientos U-A ó C-G. Sin embargo, el apareamiento entre la tercera posición del codón y la primera del anticodón tiene cierta flexibilidad ya que varios apareamientos de bases, fuera del estándar, son posibles. Es raro encontrar adenina (A) en la posición *wobble* del anticodón en el RNAt, ya que se encuentra convertida a inosina (I), la cual puede aparearse con A, C y U (Fig. 7, Tabla 4). Uridina (U) en la primera posición del anticodón se aparean específicamente con A, pero puede también reconocer G. C1 del anticodón reconoce G3 en el codón, mientras que G en la posición *wobble* reconoce específicamente C y alternativamente U en el codón (Agris, P. *et al*, 2007 y 2004). Otras modificaciones encontradas en el anticodón se muestran en la Tabla 4.

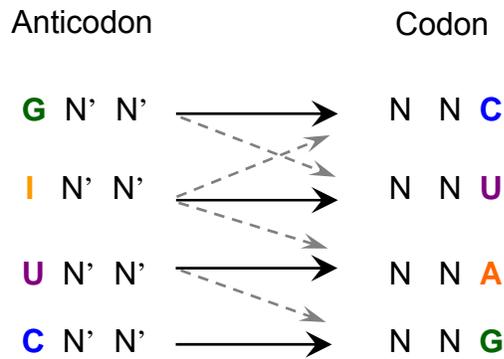


Figura 7. Apareamientos codón : anticodón complementarios (líneas continuas) y *wobble* (línea punteada gris) en eucariontes. El anticodón **G** N'N' puede reconocer los codones **NNC** y **NNT**, es decir G puede aparearse con C y T (U). La adenina en primera posición del anticodón es generalmente modificada a inosina (tomado y modificado de Percudani, R. 2001)

Tabla 4. Hipótesis *Wobble*: apareamientos posibles codón : anticodón

Hipótesis <i>Wobble</i>	
Anticodón	Codon
1ª posición	3era posición
G	C, U
C	G
I	U, C, A
xm ⁵ U ^a	G
s ² U ^b	A, G
Xo ⁵ U ^c	A, G, U

- a) Todos los derivados de 5-metileneuridina
- b) Todos los derivados de 2-tiouridina
- c) Todos los derivados de 5-oxiuridina

1.5.3. Uso de codones

Los codones sinónimos son aquellos que codifican para el mismo aminoácido. Sin embargo, entre las proteínas, no todos los codones sinónimos son usados con la misma frecuencia aun cuando son equivalentes en términos de la estructura final de la proteína. Los codones sinónimos son utilizados de manera diferente entre diferentes organismos y cada genoma tiene una estrategia de codificación particular. Por ejemplo, en *E. coli* seis codones diferentes son posibles para el aminoácido arginina (Arg), de los cuales dos parecen ser los codones óptimos: CGU, CGC. La elección de codones específicos para cada organismo está relacionada con la población de sus RNAt isoaceptores. Esta tendencia al uso de codones ocurre en la mayoría de las especies y se

cree que conduce a una co-adaptación del uso de codones y el contenido de RNAt para optimizar la eficiencia del proceso de traducción (Ikemura, T., 1985).

1.5.4. RNAt isodecodificadores.

Goodenbour y Pan (2006) describieron genes de RNAt que poseen el mismo anticodón pero muestran diferencias en el resto de la secuencia del RNAt, los RNAt isodecodificadores. Entre los genes de un isoceptor particular, los RNAt idénticos con mayor número de copias génicas corresponden al isoceptor principal. Los isodecodificadores son aquellos que tienen secuencias diferentes al isoceptor principal. Por ejemplo, entre el RNAt-Arg (ACG) en humano pueden distinguirse dos secuencias diferentes con el mismo anticodón, una con 4 copias génicas y otra con 3. La secuencia con 4 copias se asigna como el isoceptor principal y a la secuencia de 3 como el isodecodificador. Así, el RNAt-Arg (ACG) de humano tiene 1 isodecodificador (Fig. 8).

Entre eucariontes, el número de isoceptores es muy similar (de 41 a 55); sin embargo, el número de isodecodificadores varía ampliamente, pues va desde 10 en *S. cerevisiae* hasta 246 en chimpancé. Por otra parte, muchos de los cambios en la secuencia de los isodecodificadores con respecto al isoceptor principal se encuentran dentro de las regiones promotoras (Fig. 8), lo que sugiere la existencia de un control diferencial para la expresión de los genes isodecodificadores. En este sentido, se ha demostrado la expresión diferencial de tres RNAt isodecodificadores en seis tejidos humanos (Goodenbour y Pan, 2006).

	CAJA A	ANTICODON	CAJA B	
Hs . Arg . ACG1	GGGCCAG <u>TGGCGCAATGG</u>	GATAACGCGTCTGACT	ACGGATCAGAAGATTCT <u>AGGTTCGACTCCT</u>	TGGCTGGCTCG 73
Hs . Arg . ACG2	GGGCCAG <u>TGGCGCAATGG</u>	GATAACGCGTCTGACT	ACGGATCAGAAGATTCC <u>AGGTTCGACTCCT</u>	TGGCTGGCTCG 73

Figura 8. Gen de RNAt de Arg (ACG) y el RNAt isodecodificador. Estas dos secuencias pueden observarse en el genoma humano para los 7 genes de RNAt-Arg (ACG). Cuatro genes presentan la secuencia Hs.Arg.ACG1 mientras que Hs.Arg.ACG2 corresponde a 3 genes, el isodecodificador. La única diferencia, en sombreado, es el nt 50, T en la primera secuencia, C en la segunda. En negritas y subrayado, la posición de las cajas promotoras A y B. * significa que las secuencias poseen el mismo nt en esa misma posición. Secuencias de los genes tomadas de <http://lowelab.ucsc.edu/GRNAtdb/>

1.5.5. RNAt-Met de iniciación y de elongación

La iniciación de la síntesis de proteínas ocurre casi universalmente con el aminoácido metionina (Met). Existen dos clases de RNAt-Met: el RNAt-iniciador (RNAt-i), usado únicamente en el inicio de la traducción, y el RNAt-elongador (RNAt-e), que se inserta a lo largo de la cadena polipeptídica (Stortchevoi *et al*, 2003). El RNAt-i se usa exclusivamente en el inicio de la traducción al unirse directamente al sitio P del ribosoma y forma un complejo con el factor de iniciación eIF2, mientras que el elongador se une primero al sitio A y después es translocado al sitio P gracias a su unión a los factores de elongación eEF1 y GTP (Drabkin *et al*, 1998).

Las diferencias estructurales entre el iniciador y el elongador permiten que la maquinaria de síntesis de proteínas pueda discernir entre ambos tipos de RNAt. Las diferencias más importantes se mencionan en la Tabla 5 y Fig. 9.

Tabla 5. Principales determinantes para la función de elongador o iniciador.

Met iniciador	Met elongador
A1:U72	G1:C72
3 pares G:C al final del brazo del anticodón altamente conservadas entre eucariontes y bacterias	No presentan G:C
A54 y A60 (asa T Ψ C)	T54 y C/T60 (asa T Ψ C)

Estas posiciones permiten al RNAt-i o RNAt-e unirse o ser reconocidos de manera específica por sus respectivos factores de iniciación o elongación durante la síntesis de proteínas, es decir, se asegura que Met-i sólo se utilizará al principio de un polipéptido que se está sintetizando y, por su parte, el elongador se utilizará sólo a lo largo de la cadena proteica. Se ha propuesto también que algunas de estas secuencias confieren al elongador y al iniciador conformaciones distintas que permitan sus actividades específicas o su unión a los factores mencionados. Por su parte, los 3 pares G:C parecen participar en la unión del RNAt-i a la subunidad 40S (Drabkin *et al*, 1998; Kapp *et al*, 2006).

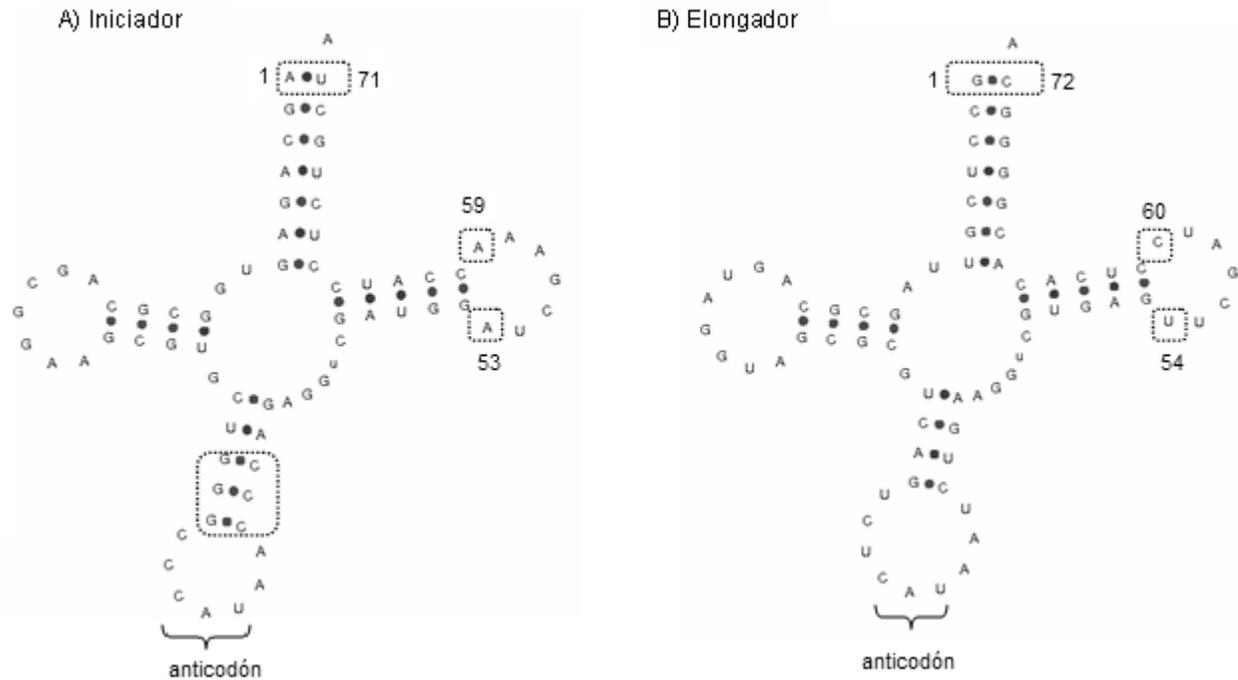


Figura 9. Principales características entre el A) RNAt iniciador y B) de elongación en humano. Las diferencias principales entre ambos RNAt se encuentran señaladas en punteado: el par de bases A1:U71 del iniciador, A1:U72 en el elongador, las 3 pb G:C en el tallo del brazo del anticodón, característica distintiva del iniciador y A53, A59 en el iniciador en contraste con las posiciones homólogas U54 y C60 del elongador. Ver texto para más detalles. Estructuras Secundarias tomadas de <http://lowelab.ucsc.edu/GRNAtdb/>

1.5.6. Distribución genómica

En eucariontes, los DNAt son genes que se transcriben de manera independiente y pueden encontrarse como genes aislados o agrupados en *clusters*. En algunos organismos, los *clusters* de DNAt se encuentran dentro de los centrómeros, como en *S. pombe* y *Caenorhabditis elegans*. Por el contrario, en organismos como *S. cerevisiae*, los DNAt se encuentran dispersos en el genoma. La mayoría de los RNAt en *S. cerevisiae* son nucleolares, al igual que algunas de las enzimas involucradas en la maduración del pre-RNAt (la RNAsa P, involucrada en la maduración del extremo 5' y la enzima Mod5 que modifica A37) (Haeusler y Engelke, 2006).

1.6. RNAt Selenocisteína

A mediados de 1980 se descubrió el aminoácido número 21, selenocisteína (Sec), codificado por el codón UGA que en el código genético universal es una señal de término para la síntesis de proteínas. Selenocisteína es un análogo de cisteína con un átomo de selenio reemplazando el grupo tiol (-SH) en dicho aminoácido (figura 10) (Allmang y Krol, 2006).

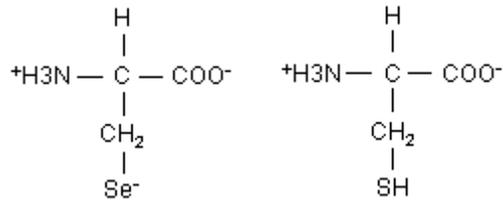


Figura 10. Estructura del aminoácido selenocisteína (izquierda), comparada con su análogo cisteína (derecha). El átomo de Se ha sustituido el grupo tiol en la cisteína.

La mayoría de las selenoproteínas caracterizadas son enzimas involucradas en reacciones redox, donde el residuo Sec es esencial para la actividad catalítica. Las selenoproteínas son esenciales en eucariontes, ya que se ha demostrado que al deletar el gen de RNAt-Sec en ratón se causa la muerte prematura del embrión, subrayando la importancia de las selenoproteínas durante el desarrollo (Johansson *et al*, 2005). En humano se han identificado ~ 25 genes codificantes de selenoproteínas, muchas con función desconocida (Kryukov *et al*, 2003). Las selenoproteínas se han encontrado en eucariontes, eubacterias y arqueobacterias; sin embargo, algunos organismos como levadura o plantas superiores no las presentan (Johansson *et al*, 2005).

Sec es incorporada a los polipéptidos durante la traducción en respuesta a los codones UGA cuando la secuencia de inserción Sec (SECIS, *selenocysteine insertion sequence*, una estructura tallo-asa específica) se encuentra en la región 3' no traducida (3'-UTR) en el RNAm en eucariontes. Los factores en *trans* involucrados en la biosíntesis e inserción de Sec son: i) gen *selC*, codifica para el RNAt-Sec, ii) *selB*, un factor de elongación Sec-específico, iii) *selD*, una selenofosfatotetasa (SPS), iv) *selA*, Sec sintasa y v) una proteína de unión a SECIS (Johansson *et al*, 2005). Selenocisteína no se presenta como un aminoácido libre, la primera etapa de su biosíntesis consiste en cargar una serina en el RNAt específico para Sec, posteriormente, el Ser-RNAt-Sec es convertido a Sec-RNAt-Sec por una selenocisteína sintasa (Allmang y Krol, 2006).

1.6.1. Estructura y función del RNAt-Sec

La estructura secundaria del RNAt-Sec posee características específicas que lo distinguen de un RNAt convencional pues son considerablemente más largos que el resto de los RNAt (en algunas bacterias llegan a ser de 100 pb). Hay dos características principales que los distinguen de un RNAt clásico: 1) poseen un brazo D de 6 pb, en vez de 3-4 pb en el resto de los RNAt. Además, el asa D se encuentra más cerrada que en el resto de los RNAt, al poseer únicamente 4 nucleótidos. 2) la secuencia desde el brazo aceptor al brazo T es más larga en RNAt-Sec (13 pb) que en los RNAt clásicos (12 pb). Estas características parecen ser señales para la unión a factores de elongación específicos (SelB en bacterias) u otro tipo de proteínas. Por su parte, los brazos D y T parecen ser determinantes para la biosíntesis de Sec en eucariontes (Allmang y Krol, 2006). Además de los brazos D y aceptor extendidos, el RNAt-Sec presenta un brazo variable (V) atípicamente más largo (Fig. 11)

A nivel de modificaciones post-transcripcionales, el RNAt-Sec contiene sólo 4 bases modificadas (m^1A58 , $\psi55$; i^6A37 y dos modificaciones posibles en U34), muy pocas comparando con los RNAt convencionales que poseen 15-17 (Fig. 11) (Allmang y Krol, 2006; Hatfield y Gladyshev, 2002).

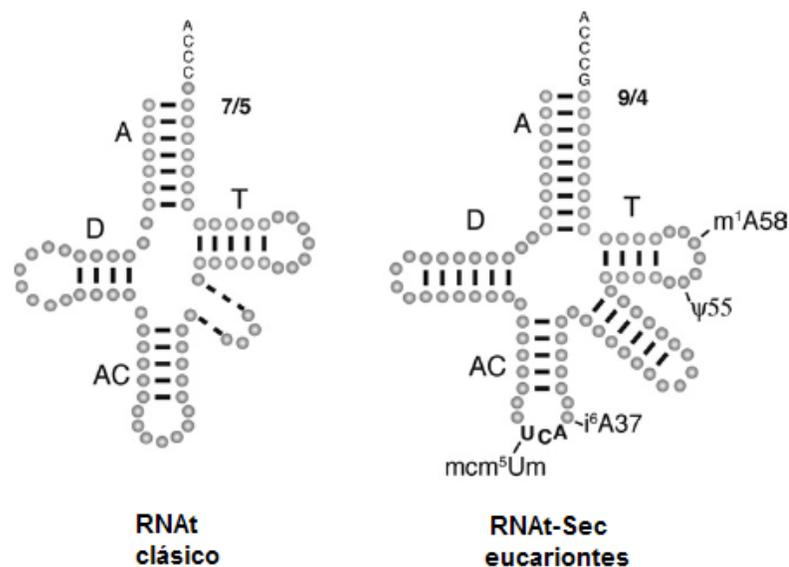


Figura 11. Comparación de estructura secundaria de un RNAt convencional y el RNAt-Sec. A, brazo del aminoácido; D, brazo D; AC, brazo del anticodón; T, brazo T. 7/5 y 9/4 indican el número de pares de bases que van del brazo del aminoácido al brazo T. El brazo D es más largo y el asa D es más cerrada en el RNAt-Sec que en los RNAt clásicos. El brazo V en el RNAt clásico (línea punteada) es de longitud variable, mientras que en el RNAt-Sec siempre es largo (~ 7 nt). Se señalan en el RNAt-Sec las modificaciones post-transcripcionales posibles (i^6A37 , isopenteniladenosina, ver texto) y se omiten en el RNAt clásico. Tomado de Allmang y Krol, 2006.

Se ha reportado que el DNAt-Sec se presenta como un gen de copia única en los genomas de los vertebrados secuenciados hasta el momento, siendo el pez cebra la única excepción conocida hasta el momento con dos copias del gen. En los genomas de humano y conejo se ha encontrado un pseudogen y tres han sido identificados en hámster (Hatfield y Gladyshev, 2002).

1.6.2. Transcripción y promotores del RNAt-Sec.

La transcripción del gen de DNAt-Sec es compleja y, en contraste con los RNAt convencionales, regulada por una combinación de elementos tanto extragénicos como intragénicos. El RNAt-Sec se distingue además porque a) el SIT coincide con el extremo 5' del transcrito, mientras que otros RNAt tienen una secuencia líder 5' que debe ser procesada y b) la caja A se encuentra modificada por la inserción de dos nt (Hatfield y Gladyshev, 2002).

El promotor basal del RNAt-Sec consiste en cuatro elementos: 3 elementos externos y uno interno. El elemento proximal (PSE), caja TATA y el elemento activador AE se encuentran río arriba del gen, de manera similar a los promotores de los genes RNAsn U6 y 7SK. El cuarto elemento es la caja B, localizada dentro de la secuencia codificadora y uno de los dos elementos promotores encontrados en los RNAt clásicos (Figuras 12 y 2).

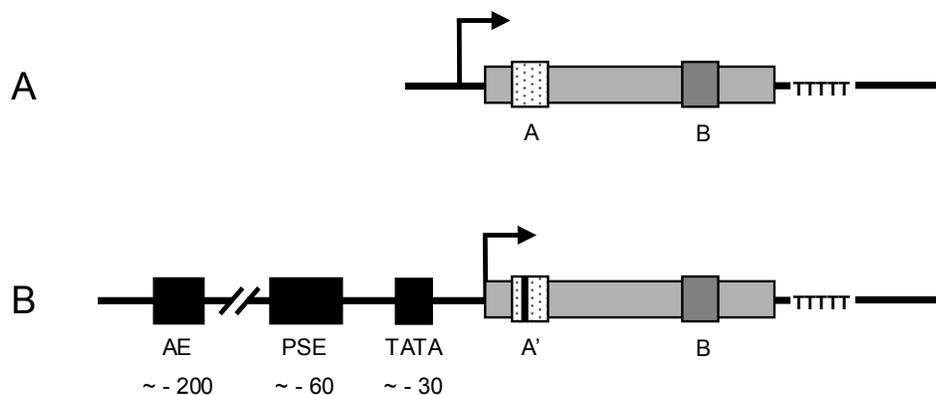


Figura 12. Elementos promotores en los genes de A) RNAt convencionales, se observa la caja A y la caja B intragénicas, y B) RNAt-Sec, que presentan los elementos intragénicos (la caja A modificada y caja B) y los extragénicos (caja TATA, PSE, AE). Los genes de RNAt en verde, la flecha indica el SIT y se muestra la secuencia de Ts en el extremo 3'.

Como se recordará, las cajas A y B de los RNAt son suficientes para activar la transcripción (Fig. 12A); sin embargo, en los RNAt-Sec no activan la transcripción por sí solas. En cambio, las cajas A y B del RNAt-Sec parecen ser necesarias para el correcto inicio de la transcripción y/o para el procesamiento adecuado de los transcritos (Park *et al*, 1995 y 1996). La caja A se encuentra modificada en los RNAt-Sec por la inserción de dos nt y parece no participar activamente en la transcripción del gen. Por su parte, la caja B no es esencial *in vivo* pero parece tener un efecto estimulante en la transcripción (Carbon y Krol, 1991; Park *et al*, 1995 y 1996).

La caja TATA es esencial para la transcripción de los genes de RNAt-Sec. La proteína TBP (*TATA binding protein*) parece unirse a esta caja TATA, lo que sería poco usual ya que en el resto de los RNAt (sin caja TATA), TBP participa en el complejo de transcripción a través de interacciones proteína-proteína (Hubert *et al*, 1996; Myslinski *et al*, 1993). El elemento PSE va aproximadamente de -72 a -49 río arriba del gen de DNAt-Sec y se requiere para una transcripción eficiente del gen, pues *in vivo*, las mutaciones en este elemento PSE impiden por completo la transcripción mientras que *in vitro* la reducen pero no por completo (Carbon y Krol, 1991; Park *et al*, 1995). Tanto la caja TATA como el elemento PSE son elementos esenciales de los promotores de los genes de RNAsn U6 en vertebrados. El elemento activador (AE, *activator element*) se requiere en tRNA-Sec para estimular la transcripción del promotor basal y consta de 15 pb entre -209 y -195 (Fig. 12B) (Myslinski *et al*, 1991). Ya se ha caracterizado el factor Staf (*selenocysteine tRNA gene transcription activating factor*), el cual se une al AE para mediar sus propiedades de activación (Schuster *et al*, 1995)

2. ANTECEDENTES

En tripanosomátidos, todos los RNAsn son sintetizados por Pol III, así como los RNAt, el RNAr 5S y el RNA 7SL. En cinetoplástidos, ya se ha identificado las cajas A y C del RNAr 5S (Cordingley, 1985; Hernández-Rivas *et al*, 1992), aunque ninguno de estos elementos ha sido analizado funcionalmente.

Acerca de los RNAt, en *T. brucei* se han identificado previamente muchos de los RNAt isoaceptores presentes en su genoma (Tan *et al*, 2002). En cuanto a su distribución, se ha descrito que se encuentran en *clusters* y que se observan secuencias intergénicas cortas entre los genes de RNAt (Mottram, *et al*, 1991; Tan *et al*, 2002). Asimismo, se ha descrito que los genes de RNAt en tripanosomátidos poseen las cajas A y B clásicas, y que estos elementos se requieren para su expresión *in vivo* (Nakaar *et al*, 1994). Resulta interesante el hecho de que en tripanosomátidos los genes de RNAsn poseen un gen de RNAt orientado de manera divergente (en la hebra contraria) y las cajas A y B del gen de RNAt son esenciales para la expresión de los RNAsn (Hernández-Rivas *et al*, 2007).

En tripanosomátidos, como en otros organismos, una secuencia de Ts sirve también como señal para el término de la transcripción. Por ejemplo, la transcripción en *L. major* del RNAt en el cromosoma 3 y en *T. brucei* del RNAsn U2 termina en una secuencia de varias Ts. Hasta la fecha, no se han encontrado proteínas involucradas en la terminación de la transcripción por Pol III en tripanosomátidos (Hernández-Rivas *et al*, 2007).

La reciente publicación de los genomas completos de los *Trypanosoma* ha permitido la realización de varios tipos de análisis a nivel genómico. Sin embargo, éstos se han enfocado únicamente en genes transcritos por Pol II. Si bien los RNAt y RNAr 5S están anotados en las bases de datos, un análisis detallado de sus secuencias y su organización genómica no se ha llevado a cabo a la fecha; menos aún, un análisis comparativo entre *Trypanosoma* y otros organismos. Se espera que el análisis bioinformático de las secuencias de genes de RNAt y 5S RNAr revele algunas características particulares de los tripanosomátidos. Las predicciones más interesantes que arroje el presente trabajo serán validadas experimentalmente en un futuro cercano.

3. OBJETIVOS

3.1. OBJETIVO GENERAL

Analizar *in silico* la organización y estructura de los genes del RNAt y del RNAr 5S y en el genoma de *Leishmania major*, *Trypanosoma brucei* y *Trypanosoma cruzi*.

3.2. OBJETIVOS ESPECÍFICOS

1. Determinar el número y localización de los genes de RNAt y RNAr 5S en el genoma de *L. major*, *T. brucei* y *T. cruzi*.
2. Construir mapas de los cromosomas con los genes RNAt y RNAr 5S correspondientes y determinar una posible relación estructural y/o funcional con genes transcritos por Pol II.
3. Identificar los probables promotores internos: cajas A y B para genes de RNAt; y cajas A y C y el Elemento Intermedio en los genes de RNAr 5S. Obtener las secuencias consenso para cada uno de estos elementos.
4. Generar las estructuras secundarias de los RNAt y los RNAr 5S.
5. Analizar el número de RNAt isoaceptores presentes en los genomas de *L. major*, *T. brucei* y *T. cruzi*.
6. Determinar la presencia de genes de RNAt isodecodificadores en el genoma de *L. major*, *T. brucei* y *T. cruzi*.
7. Determinar el porcentaje de uso de codones en los genomas de *L. major*, *T. brucei* y *T. cruzi*.
8. Con base en el uso de codones, determinar si existe correlación estadística entre el número de genes de RNAt y el uso de codones en *L. major*, *T. brucei* y *T. cruzi*.

4. METODOS

Los genomas de *L. major*, *T. brucei* y *T. cruzi* se encuentran disponibles en las bases de datos de GeneDB (<http://www.genedb.org/>), que fueron analizadas *in silico* para buscar todos los genes de RNAr 5S y RNAt de cada especie. Las bases de datos fueron analizadas con el programa tRNAscan-SE (<http://lowelab.ucsc.edu/tRNAscan-SE/>) (Lowe y Eddy, 1997) para verificar la presencia y la identidad de los genes de RNAt, así como para obtener las estructuras secundarias. El programa analiza la presencia de 4 características en la secuencia a analizar: las secuencias de las cajas A y B, la distancia entre estos dos elementos y la distancia entre la caja B y la señal de término de Pol III. Así, este programa detecta el RNAt en la secuencia analizada y predice el anticodón, y los intrones (de estar presentes), arrojando una calificación (un *score* >30 indica que la estructura es un RNAt clásico). tRNAscan puede también reconocer pseudogenes y estructuras *tRNA-like* (secuencias con un *score* bajo) (Lowe y Eddy, 1997). Los intrones en los genes de RNAt-Tyr fueron localizados con este programa. Una vez verificados los RNAt e identificada su localización, se construyeron los mapas de organización de estos genes para *L. major* y *T. brucei*. La información sobre la relación espacial entre los genes de Pol III y Pol II y los mapas de sintenia se obtuvo de GeneDB.

Las probables regiones de control interno de los diferentes genes fueron identificadas y comparadas usando el programa ClustalW (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>). Una vez alineadas las secuencias de las cajas, pudimos realizar el análisis para obtener las secuencias consenso.

Los genes de RNAt-Sec (Cassago, *et al*, 2006) y el gen de sRNA76 (Béjà, *et al*, 1993) se identificaron por búsqueda tipo BLAST utilizando las secuencias correspondientes reportadas en la literatura. Es importante mencionar que el tRNA-Sec no es fácilmente detectado por tRNAscan debido a su estructura secundaria atípica, por lo que optamos por realizar la búsqueda tipo BLAST.

Los RNAt isodecodificadores fueron identificados utilizando las secuencias de los genes de RNAt isoaceptores obtenidas de GeneDB y alineándolas en el sitio ClustalW para conocer el porcentaje de identidad.

Los datos sobre el uso de codones se obtuvieron de la base de datos del Instituto Sanger (www.sanger.ac.uk) para *L. major* (8272 genes de proteínas) y de la página web Kasuza (www.kazusa.jp) para *T. brucei* (5140 genes de proteínas). Los datos de uso de codones para *T. cruzi* se obtuvieron utilizando 1779 genes de proteínas tomados al azar a lo largo del genoma (del sitio del Instituto Sanger); los cálculos se realizaron en la página: <http://www.bioinformatics.org/sms2/>. El coeficiente de Spearman y los cálculos estadísticos sobre la secuencia de Ts se obtuvieron con Excel y el programa GraphPad Prism 5 (<http://www.graphpad.com>).

En el caso del RNAr 5S, las estructuras secundarias se generaron utilizando RNAfolding (<http://mfold.burnet.edu.au/>). Para obtener las secuencias consenso de las cajas promotoras se utilizaron todos los genes de RNAr 5S anotados en GeneDB de *L. major* y *T. brucei* y 10 genes tomados al azar de *T. cruzi*. Las secuencias de los probables elementos promotores fueron alineadas con ClustalW y posteriormente analizadas para obtener una secuencia consenso.

5. RESULTADOS

5.1. Número de genes de RNA de transferencia en tripanosomátidos

El análisis de las bases de datos de GeneDB de *L. major*, *T. brucei* y *T. cruzi* indica la presencia de 82, 65 y 115 genes de tRNA, respectivamente. Como primer punto de nuestro análisis, confirmamos la identidad de los 82 genes anotados en el genoma de *L. major*. Sin embargo, en los genomas de *T. brucei* y *T. cruzi* se encontraron algunas discrepancias en cuanto a la identidad de algunos RNAt. Con los RNAt comprobados se construyeron mapas genómicos en *L. major* (Fig. 13) y *T. brucei* (Fig. 14). Cabe mencionar que el genoma secuenciado de *T. cruzi* no se encuentra ensamblado en cromosomas y la información se muestra por *contigs* o fragmentos de DNA secuenciados. Por esta razón no se pudieron generar los mapas genómicos de los RNAt en *T. cruzi*.

En *T. brucei* se encontraron tres genes de RNAt mal anotados, dos de ellos localizados en el cromosoma 8, en el *cluster* V (Fig. 14): Tb08_tRNA_Leu_2 (Tb927.8.6564) y Tb08_tRNA_Gln_2 (Tb927.8.6566). En GeneDB, estos dos genes se encuentran anotados en la cadena sentido, pero el análisis de la secuencia reveló que en realidad se encuentran en la hebra antisentido (Fig. 14). El tercer caso fue el del gen Tb04_tRNA_Sec_1 (Tb927.4.4663) que se encuentra anotado como un gen de selenocisteína en el cromosoma 4; sin embargo, nuestro análisis indicó que en realidad este gen corresponde al sRNA76 (Fig. 14, *cluster* chr04-V), una secuencia tipo RNAt que se asocia con el RNA 7SL en tripanosomátidos (Béjà *et al*, 1993). Con el objeto de localizar el gen auténtico de RNAt-Sec en *T. brucei*, se llevó a cabo un análisis tipo BLAST del genoma de dicho parásito, usando como blanco la secuencia del gen de RNAt-Sec previamente reportada (Cassago *et al*, 2006). De esta manera, fueron identificados dos genes de RNAt-Sec en el cromosoma 9 de *T. brucei* (Fig. 14, *clusters* chr09-II y III).

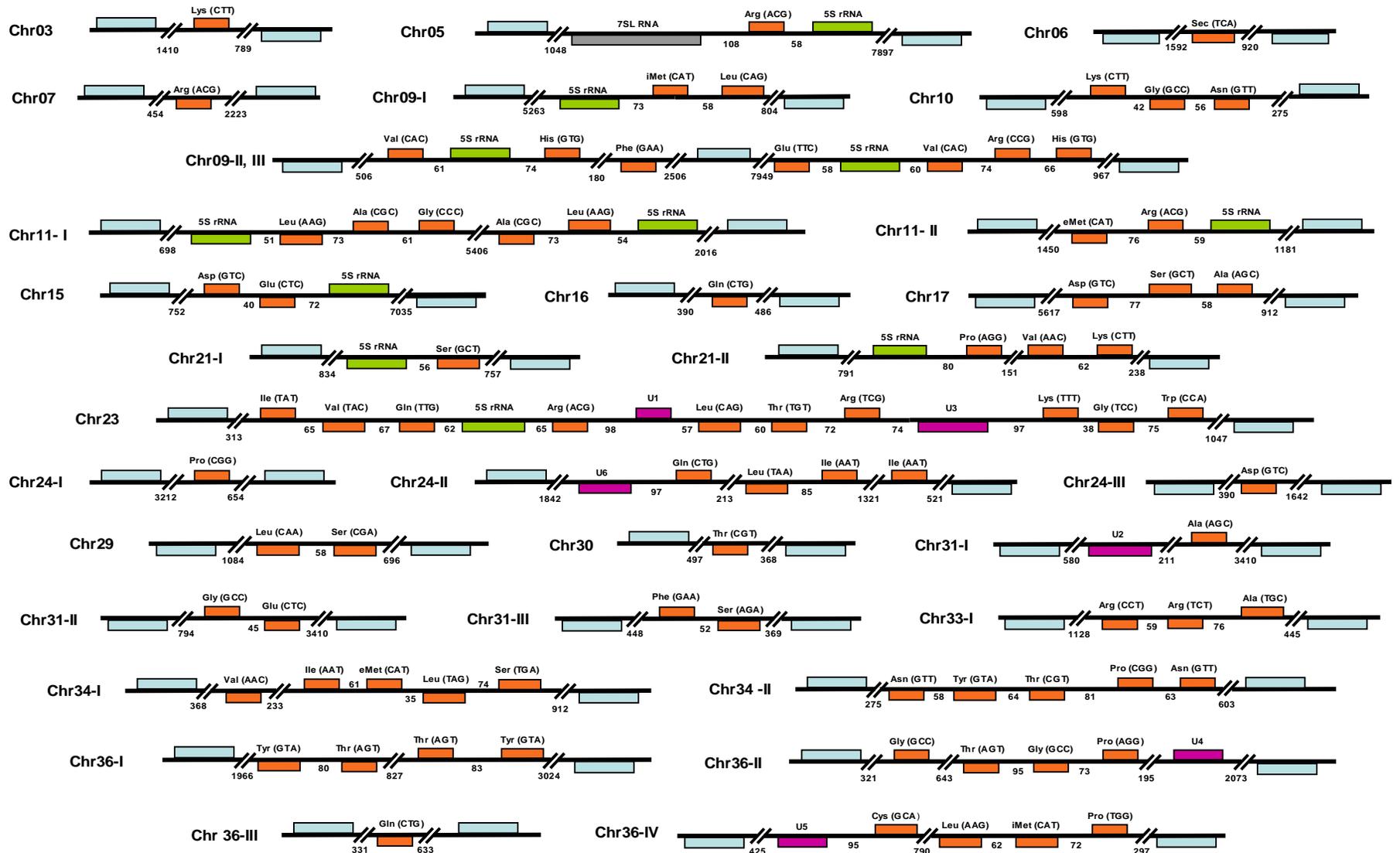


Figura 13. Organización de los genes de RNAt en *L. major*. En naranja se muestran los 83 genes encontrados en el genoma del parásito y los anticodones se indican entre paréntesis. Los genes de RNAr 5S, RNAsn y RNA 7SL se presentan en verde, violeta y gris, respectivamente. El gen de RNAt-Sec se encuentra en el *cluster* Chr06 (localización 69,586 a 69,673). Los genes están a escala y se indican el tamaño de regiones intergénicas en pb. Los genes codificadores de proteínas que flanquean los genes transcritos por Pol III se muestran en azul (no están a escala). Los genes codificadores de proteínas se utilizaron como delimitantes de los locus de Pol III, por lo que los *clusters* Chr09-II y III se consideran dos *loci* de Pol III independientes.

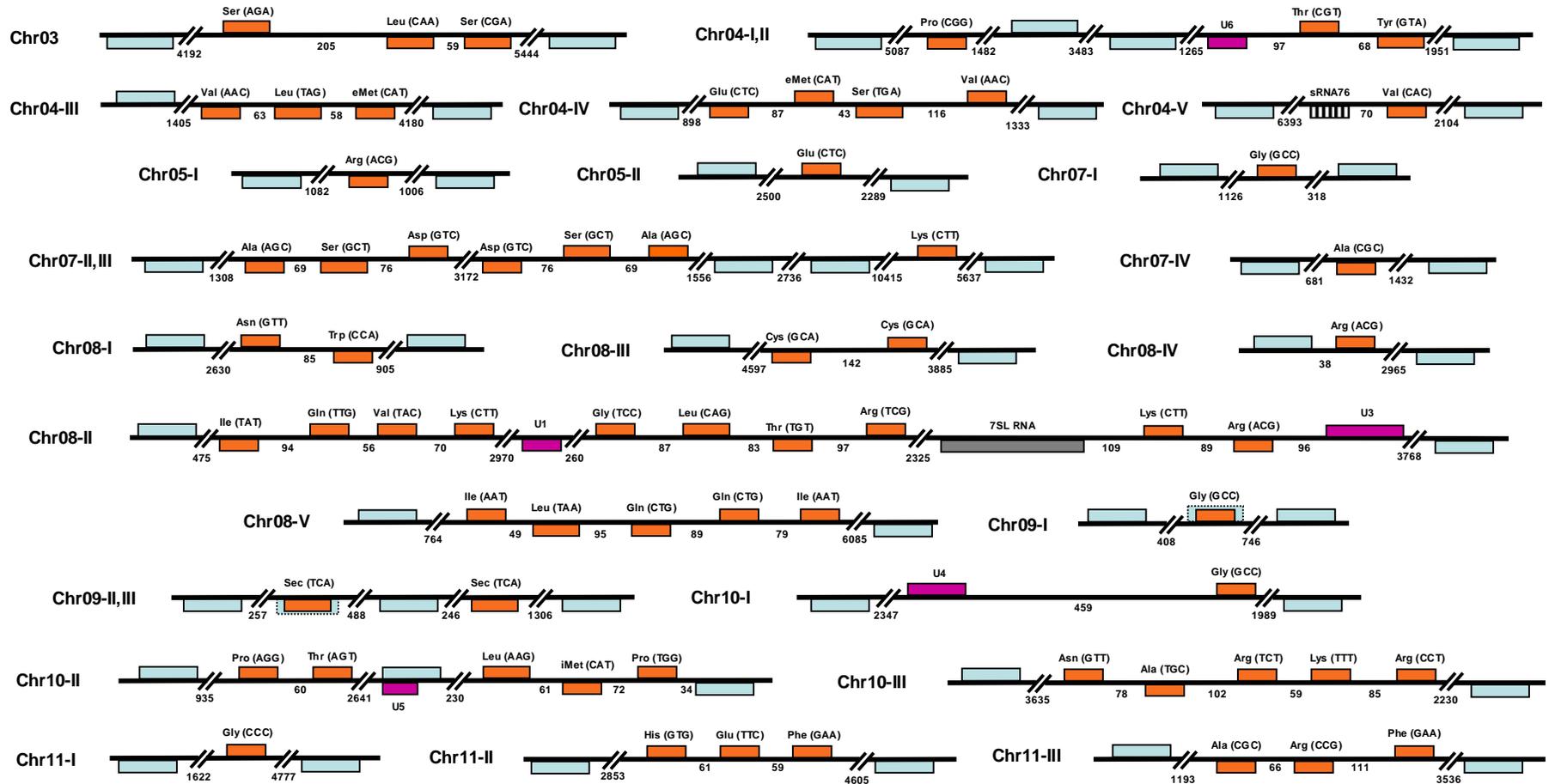


Figura 14. Organización de los genes de RNAt en *T. brucei*. En naranja se muestran los 66 genes encontrados en el genoma del parásito y los anticodones se indican entre paréntesis. Los genes de RNAsn, RNA 7SL y el sRNA76 se presentan en violeta, gris y a rayas, respectivamente. El *cluster* Chr09-II se sobrelapa con un gen putativo de Pol III (Tb09.160.1080, en azul con marco punteado), anotado como una proteína putativa y lo mismo sucede en el *cluster* Chr09-I, donde el RNAt sobrelapa con la proteína hipotética Tb09.211.4080. El gen RNAsn U5 (Chr10-II) también se sobrelapa con un gen codificador de proteína (Tb10.6k15.2990), pero localizado en la hebra contraria. Los genes transcritos por Pol III están a escala y se indican el tamaño de regiones integénicas en pb. Los genes codificadores de proteínas se muestran en azul (no están a escala). En el locus Chr08-V, los genes de RNAt Leu-TAA (Tb_08_TRNA_Leu_2) y el primero de Gln-CTG (Tb_08_TRNA_Gln_2) están mal anotados en GeneDB respecto a la hebra en la que se encuentran, y lo mismo para el gen del RNAsn U6, en el *locus* Chr04-II. Los dos genes de RNAt-Sec fueron localizados en los *clusters* Chr09-II y III (en las posiciones 505,606 a 505,693 y 509,443 a 509,530).

Un análisis similar llevado a cabo en el genoma de *L. major* nos permitió localizar un gen de RNAt-Sec en el cromosoma 6 (Fig. 13, *cluster chr06*). En *T. cruzi*, se localizaron ocho copias del gen de RNAt-Sec dispersas en diferentes *contigs* a lo largo del genoma, organizados como genes independientes (Tabla 6).

Tabla 6. Copias del RNAt-Sec en *T. cruzi*.

<i>Contig</i>	Localización	Hebra DNA
18653	556-643	Sentido
11959	770-857	Antisentido
14021	87-174	Antisentido
28162	15-102	Sentido
11097	441-528	Sentido
6957	12609-12696	Antisentido
6957	9112-9199	Antisentido
6536	26013-26100	Antisentido

En *T. cruzi* se encontraron cuatro genes mal anotados en las bases de datos: tres de ellos anotados como genes de RNAt correspondientes a Val-CAC: Tc00.1047053457717.10, Tc00.1047053483321.10 y Tc00.1047053506321.220. Sin embargo, no parecen corresponder al aminoácido asignado (análisis realizado con tRNAscan-SE), pues sólo presentan 20%, 20% y 61% de identidad con el gen real de RNAt-Val (CAC) (Tc00.1047053506459.249) (Fig. 15A) que consideramos como el real, ya que presenta 98% y 100% de identidad con los genes homólogos de *L. major* y *T. brucei*, respectivamente (Fig. 15E). La secuencia de los genes Tc00.1047053457717.10, Tc00.1047053483321.10 y Tc00.1047053506321.220 difiere de la del gen “real” de valina en varios aspectos: en ellos no se encuentra el anticodón CAC en la posición esperada y en su estructura secundaria se pueden observar inserciones en el asa del anticodón, que es de 9 nt en vez de 7 nt en el gen real de RNAt-Val; así como un plegamiento diferente de la molécula en la región del brazo D (Fig. 15B-15D).

El cuarto gen mal anotado es Tc00.1047053507579.16, señalado como correspondiente a Ile-TAT; sin embargo, sus cajas promotoras se encuentran en posición diferente a lo típico y no posee un tracto de Ts en su extremo 3' (Fig. 16A); además, muestra sólo 22% de identidad con los otros dos genes anotados de Ile-TAT (Fig. 16B) y su estructura secundaria no se pliega de manera típica, con los brazos T y del aminoácido más cortos y el asa del anticodón con inserciones (9 nt en vez de 7 pb en el gen Ile-TAT real) (Fig. 16D-E). Por esto, consideramos que Tc00.1047053507579.16 podría ser un pseudogen o un RNAt-like. Los genes que consideramos "reales" de Ile-TAT en *T. cruzi* (Tc00.1047053504427.231 y Tc00.1047053508043.11) se parecen 91% y 94% con los genes de *L. major* y *T. brucei*, respectivamente (Fig. 16C).

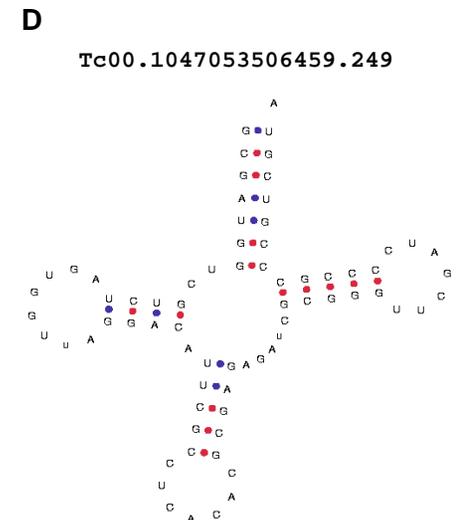
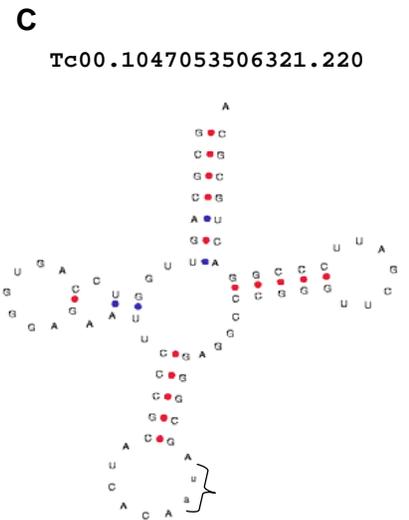
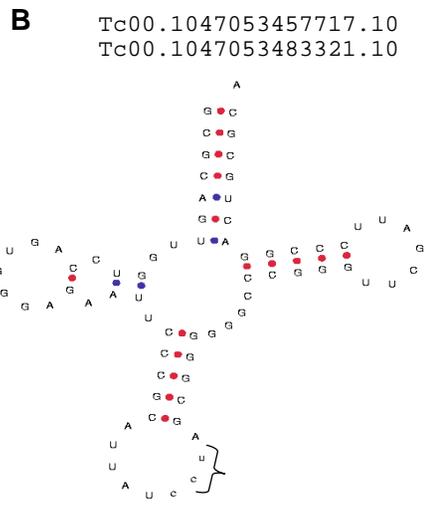
Finalmente, en *T. cruzi*, se localizó un gen de RNAt-Ala (TGC) que no se encuentra anotado en GeneDB. Este gen se localizó en el contig 8001 al realizar una búsqueda tipo BLAST con el gen de alanina que se encontraba anotado como gen de copia única. Una vez indentificada esta secuencia, se confirmó la presencia del RNAt con tRNAscan (Apéndice I).

A

```

Tc00.1047053457717.10    GCGCAGTTGGTCCAGTGGG-AGAATTCCCGCATTATCCTAGCGGGGGGCC-CGGGTTTCGATTCCCGGACTGCGCA 73
Tc00.1047053506321.220   GCGCAGTTGGTCCAGTGGG-AGAATTCCCGCATCACAATAGCGGGAGGCC-CGGGTTTCGATTCCCGGACTGCGCA 73
Tc00.1047053506459.249   GCGATGGTTCGTCTAGTGGTTAGGACATTCGCCTCACAC--GCGAGAGATCGCGGGTTTCGATCCCCGCCCGTCGTA 73
*** * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```



E

```

Lm.TRNAVAL                GCGATGGTTCGTCTAGTGGTTAGGACATTCGCCTCACACGCGAGAGATCGCGGGTTCAATCCCCGCCCGTCGTA 73
Tb04_tRNA_Val_3           GCGATGGTTCGTCTAGTGGTTAGGACATTCGCCTCACACGCGAGAGATCGCGGGTTTCGATCCCCGCCCGTCGTA 73
Tc00.1047053506459.249    GCGATGGTTCGTCTAGTGGTTAGGACATTCGCCTCACACGCGAGAGATCGCGGGTTTCGATCCCCGCCCGTCGTA 73
*****

```

Figura 15. Análisis de los RNAt-Val putativos en *T. cruzi*. A) Alineamiento que muestra la identidad de secuencia de los genes tRNA-like [Tc00.1047053457717.10, Tc00.1047053483321.10 (cuyas secuencias son idénticas 100% mostrándose sólo la secuencia del primero) y Tc00.1047053506321.220] y el gen real de Val-CAC (Tc00.1047053506459.249), en el que se marca el anticodón CAC subrayado y en negritas. B, C y D) Estructuras secundarias de los genes de RNAt analizados en el panel A. Nótese en B y C las inserciones en el asa del anticodón (marcadas con corchete) y las diferencias en el plegamiento de la estructura secundaria respecto a D (ver texto). E) Alineamiento de las secuencias del RNAt-Val CAC entre *L. major*, *T. cruzi* y *T. brucei*. Las secuencias son idénticas entre los dos últimos y con *L. major* 98%. Lm.TRNAVAL.CAC (LmjF09.TRNAVAL.01, LmjF09.TRNAVAL.02), Tb04_tRNA_Val_3, Tc00.1047053506459.249. En sombreado se muestra el anticodón CAC, presente en los genes de las tres especies.

Así, nuestro análisis indica que *L. major* posee 83 genes (82 anotados y 1 de Sec), *T. brucei* posee 66 genes (64 anotados correctamente, más 2 de Sec) y *T. cruzi* posee 120 genes de RNAt (115 anotados, 8 copias de selenocisteína, 1 copia de alanina no anotada y 4 genes mal anotados que se eliminaron de la lista final). La información completa sobre los genes de RNAt, sus anticodones respectivos y localización se encuentra en el Apéndice I.

5.2. RNAt isoaceptores

Tras analizar y verificar la identidad de todos los genes de RNAt en tripanosomátidos, pudimos identificar los RNAt isoaceptores presentes en los genomas, es decir, aquellos RNAt que acarrean el mismo aminoácido pero poseen diferente anticodón. Aunque el número total de genes de RNAt es diferente entre las tres especies, el número de isoaceptores encontrados fue 46 en *L. major*, *T. brucei* y *T. cruzi*; 45 de los isoaceptores corresponden a los aminoácidos universales y 1 para selenocisteína (Tabla 7). Estos 46 isoaceptores son suficientes para leer los 61 codones posibles para los diferentes aminoácidos, más uno del aminoácido selenocisteína. Aunque algunos autores consideran que existen dos isoaceptores para el aminoácido metionina, el iniciador y elongador, para fines prácticos nosotros los consideraremos como un solo isoaceptor.

Dieciséis anticodones no fueron encontrados en el genoma de los tripanosomátidos; sin embargo, sus codones correspondientes sí se encuentran en los genes codificadores de proteínas. Por ejemplo, el anticodón de cisteína ACA no se encuentra en ninguno de los RNAt en tripanosomátidos, pero el codón correspondiente TGT sí se encuentra en las proteínas (Tabla 7). Esto es posible porque algunos RNAt pueden aparearse con codones que sean complementarios sólo en las dos primeras posiciones del codón, permitiendo que el apareamiento entre la tercera base del codón y la primera del anticodón sea flexible (hipótesis *Wobble* o del Tambaleo).

Se observó que entre los tripanosomátidos los anticodones faltantes comenzaban con A o G (Tabla 7). Es importante mencionar que en la posición *wobble* del anticodón, en el RNAt maduro, A es siempre convertida a inosina (I), que es una purina modificada que puede aparearse con A, C y U (ver Fig. 7, en la introducción). Así, cuando para un aminoácido no se encuentra el anticodón con G en la primera posición (G1), el RNAt isoceptor con A1 (modificada a I) es el encargado de reconocer los codones con C o U en la tercera posición (Fig. 17B). Y viceversa, cuando no se encuentra el anticodón con A1 en el genoma, el RNAt isoceptor con G1 lee C3 o U3 en los codones complementarios (Fig. 17A). Así, nuestro análisis reveló que los tripanosomátidos utilizan una estrategia de codificación conocida como “recambio de A1 - G1” (Fig. 17). Se observó también que los nucleótidos de “recambio” o “repuesto” A1 y G1, son utilizados de la misma manera, pues el 50% (8/16) de los codones U3 o C3 son leídos por A1 (modificada a I) y 50% (8/16) por G1 (Tabla 7).

Tabla 7. Repertorio de genes de RNAt y uso de codones en Trityps

¹ Un isodecodificador fue identificado en *L. major*, uno en *T. brucei* y seis en *T. cruzi*. El porcentaje de identidad entre RNA isodecodificadores se muestra entre paréntesis. El resto de los RNAt con el mismo anticodón son 100% idénticos. ² Los anticodones con fondo negro son aquellos que reconocen más de un codón, por ejemplo, el anticodón de tirosina, GUA, reconoce dos codones: TAT y TAC, ya que el anticodón AUA no está presente en tripanosomátidos. A1 en el anticodón es siempre modificado a inosina (I). ³ Ala, alanina, Arg, arginina, Asn, asparagina, Asp, ácido aspártico, Cys, cisteína, Gln, glutamina, Glu, ácido glutámico, Gly, glicina, His, histidina, Ile, isoleucina, Leu, leucina, Lys, lisina, Met, metionina, Phe, fenilalanina, Pro, prolina, Ser, serina, Thr, treonina, Trp, triptófano, Tyr, tirosina, Val, valina, Sec, selenocisteína.

Amino ácido ³	Codon	DNA _t anticodon	<i>L. major</i> número de copias DNA _t	<i>T. brucei</i> número de copias DNA _t	<i>T. cruzi</i> número de copias DNA _t	Anticodón ²	<i>L. major</i> Uso de codones %	<i>T. brucei</i> Uso de codones %	<i>T. cruzi</i> Uso de codones %	
Ala	GCT	AGC	2	2	2 (98%) ¹	IGC	1.83	2.08	1.61	
	GCC	GGC	0	0	0		3.61	1.82	1.62	
	GCA	TGC	1	1	2	UGC	2.02	2.32	2.16	
	GCG	CGC	2	2 (97%) ¹	2	CGC	3.62	2.07	2.36	
Arg	CGT	ACG	4	3	4 (98%) ¹	ICG	1.05	1.59	1.60	
	CGC	GCG	0	0	0		3.18	1.45	1.75	
	CGA	TCG	1	1	2	UCG	0.73	0.9	1.58	
	CGG	CCG	1	1	2	CCG	1.38	1.21	1.85	
	AGA	TCT	1	1	2	UCU	0.27	0.68	1.56	
	AGG	CCT	1	1	2	CCU	0.55	1	1.86	
Asn	AAT	ATT	0	0	0	GUU	0.53	1.82	1.37	
	AAC	GTT	3	2	4		2.07	1.92	1.45	
Asp	GAC	GTC	3	2	2	GUC	3.42	2.27	1.60	
	GAT	ATC	0	0	0		1.44	2.81	1.57	
Cys	TGT	ACA	0	0	0	GCA	0.42	1.09	1.78	
	TGC	GCA	1	2	2		1.46	1.13	2.34	
Gln	CAA	TTG	1	1	2	UUG	0.76	1.69	1.66	
	CAG	CTG	3	2	4	CUG	3.32	2.1	1.8	
Glu	GAA	TTC	1	1	1	UUC	1.14	3.17	2.17	
	GAG	CTC	2	2	4 (95%) ¹	CUC	3.32	3.81	2.21	
Gly	GGT	ACC	0	0	0	GCC	1.25	2.27	1.51	
	GGC	GCC	4	3	4		3.36	1.49	1.95	
	GGA	TCC	1	1	2	UCC	0.64	1.56	2.16	
	GGG	CCC	1	1	2	CCC	1.19	1.39	1.64	
His	CAT	ATG	0	0	0	GUG	0.65	1.13	1.37	
	CAC	GTG	2	1	4 (98%) ¹		2.04	1.3	1.64	
Ile	ATT	AAT	3	2	4	IAU	0.82	1.9	1.42	
	ATC	GAT	0	0	0		1.88	1.16	1.10	
	ATA	TAT	1	1	2	UAU	0.27	1	0.75	
Leu	TTA	TAA	1	1	2	UAA	0.16	0.98	0.89	
	TTG	CAA	1	1	2	CAA	1.1	1.96	2.33	
	CTT	AAG	3	1	4	IAG	1.11	2.23	1.59	
	CTC	GAG	0	0	0		2.5	1.56	1.27	
	CTA	TAG	1	1	2	UAG	0.47	0.82	0.54	
Lys	CTG	CAG	2	1	2	CAG	3.83	1.85	1.95	
	AAA	TTT	1	1	2	UUU	0.54	2.06	1.94	
	AAG	CTT	3	3	4	CUU	2.78	2.66	1.88	
	Met	ATG	CAT	4	3	6	CAU	2.25	2.34	2.12
	Phe	TTT	AAA	0	0	0	GAA	1.04	2.05	2.17
TTC		GAA	2	2	4		1.91	1.59	1.49	
Pro	CCT	AGG	2	1	2	IGG	0.86	1.11	1.05	
	CCC	GGG	0	0	0		1.24	1.11	1.03	
	CCA	TGG	1	1	2	UGG	1.05	1.39	1.44	
	CCG	CGG	2	1	2	CGG	2.61	1.18	1.57	
Ser	TCT	AGA	1	1	2	IGA	1.02	1.26	1.37	
	TCC	GGA	0	0	0		1.69	1.25	1.24	
	TCA	TGA	1	1	2 (98%) ¹	UGA	0.73	1.35	1.41	
	TCG	CGA	1	1	2	CGA	2.17	1.16	1.2	
	AGT	ACT	0	0	0	GCU	0.73	1.51	1.21	
	AGC	GCT	2 (98%) ¹	2	2		2.6	1.35	1.60	
Thr	ACT	AGT	3	1	2	IGU	0.68	1.3	1.13	
	ACC	GGT	0	0	0		1.77	1.21	1.18	
	ACA	TGT	1	1	2	UGU	1.04	1.74	1.74	
	ACG	CGT	2	1	2	CGU	2.48	1.48	1.98	
Trp	TGG	CCA	1	1	2	CCA	1.07	1.09	2.33	
Tyr	TAT	ATA	0	0	0	GUA	0.4	1.13	0.84	
	TAC	GTA	3	1	2		1.99	1.41	1.04	
Val	GTT	AAC	2	2	2	IAC	0.86	2.29	1.49	
	GTC	GAC	0	0	0		1.92	1.14	1.26	
	GTA	TAC	1	1	2	UAC	0.54	1.26	0.83	
	GTG	CAC	2	1	1	CAC	3.82	2.87	2.65	
Sec	TGA	TCA	1	2	8					
TOTAL			83	66	120					

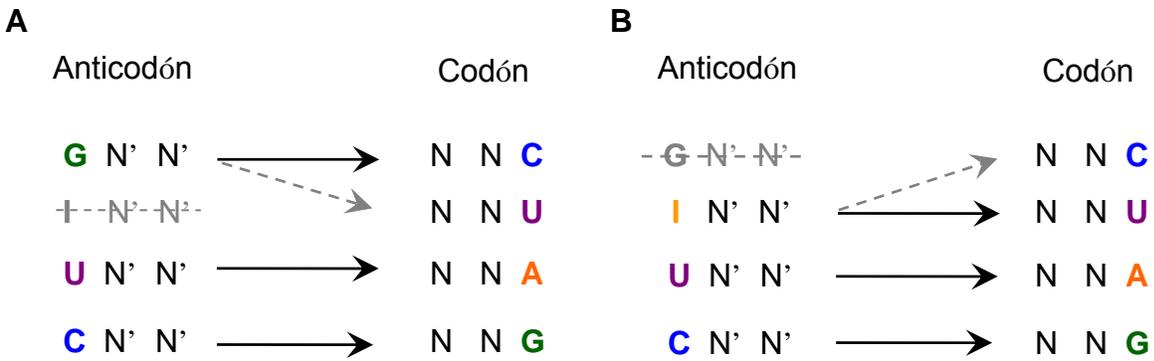


Figura 17. Estrategia de codificación utilizada por tripanosomátidos para poder leer los 61 codones del código genético. A) Cuando para cierto aminoácido no se encuentra el anticodón con A (I) en la primera posición, G lee los codones con C y U en la tercera posición del codón. B) Cuando para cierto aminoácido no se encuentra el anticodón con G en la primera posición, A (convertida a inosina, I) lee los codones con C3 y U3. Las líneas negras representan el apareamiento complementario, mientras que las líneas punteadas el apareamiento flexible o *wobble*.

5.3. RNAt de Metionina elongador e iniciador

Los genomas de *L. major*, *T. brucei* y *T. cruzi* contienen cuatro, tres y seis genes de RNAt-Met, respectivamente. Estas secuencias se analizaron y se identificaron los elongadores (Met-e) y los iniciadores (Met-i) de cada especie. Así, en *L. major* se encontraron dos iniciadores y dos elongadores, en *T. brucei* un iniciador y dos elongadores, y finalmente en *T. cruzi* dos Met-i y cuatro Met-e (Tabla 8).

Tabla 8. Lista de genes de RNAt-Met-i y RNAt-Met-e en Trityps. Se indica el nombre asignado en GeneDB

	<i>L. major</i> ¹	<i>T. brucei</i> ¹	<i>T. cruzi</i> ²
Met-i	LmjF09.TRNAMET.01 LmjF36.TRNAMET.01	Tb10_tRNA_Met_1	Tc00.1047053506251.88 Tc00.1047053508231.92
Met-e	LmjF11.TRNAMET.01 LmjF34.TRNAMET.01	Tb04_tRNA_Met_1 (Tb927.4.3199) Tb04_tRNA_Met_2 (Tb927.4.3294)	Tc00.1047053504055.87 Tc00.1047053504055.91 Tc00.1047053506435.327 Tc00.1047053506435.345
Total	4	3	6

¹ número de genes por genoma haploide

² número de genes por genoma presumiblemente diploide

Se identificaron en los tRNA^{Met-e} y tRNA^{Met-i} de los tripanosomátidos casi todas las características que los distinguen en otros organismos, y que son necesarias para que la maquinaria traduccional pueda reconocerlos. Al igual que en otros eucariontes, se encontró un apareamiento A1:U71 en los iniciadores, mientras que un par G1:C72 se localizó en los elongadores (Fig. 18A y 18B). Estas características se encontraron conservadas entre las tres especies. Otros determinantes para la función de iniciador que se encontraron conservadas entre tripanosomátidos son A53, A58 y A59 en el asa T_ψC (U54, U59 y U60 en el elongador), el par C3:G69 (G3:C70 en Met-e), las bases A19 y A20 (G19 y C20 en el elongador), un asa D y asa del anticodón formadas por 7 nt (8 y 9 nt en Met-e, respectivamente) (Fig. 18A y 18B).

Otra característica que distingue el elongador del iniciador es la presencia de tres G:C consecutivas en la parte baja del brazo del anticodón de los iniciadores. Entre los tripanosomátidos, Met-i posee estos pares G:C, pero sorpresivamente, los encontramos también en el elongador, por lo que en estos organismos aparentemente no son un determinante para distinguir entre los dos tipos de RNA-Met, como lo son en otros eucariontes (Fig. 18A y 18B).

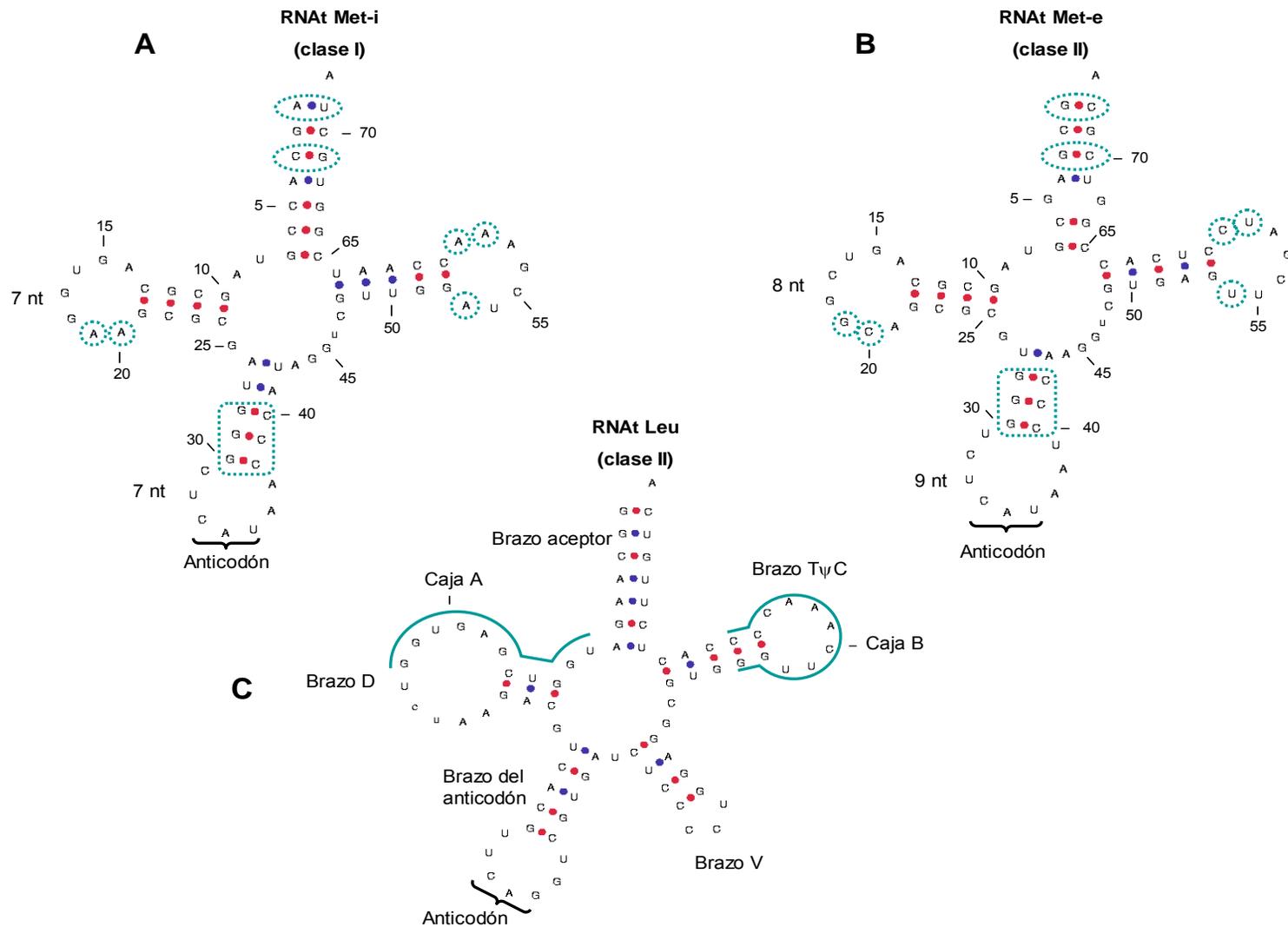


Figura 18. Estructuras secundarias del RNA^t-Met iniciador (A) y elongador (B) en *L. major*. Se muestran las características importantes para la función de Met-e y Met-i. Estas dos moléculas corresponden a moléculas de clase I. En C, se muestra la estructura secundaria de un RNA^t de clase II (Leu). Se indican los nombres de los diferentes brazos y el anticodón, así como la posición de los elementos de control interno, cajas A y B.

5.4. Organización de los genes de RNAt

En *L. major*, los 83 genes de RNAt están distribuidos en 31 *loci*, en 19 cromosomas diferentes (Fig. 13, Tabla 9). La mayoría de los genes de RNAt están organizados en *clusters* de 2 a 10 genes, los cuales pueden contener otros genes transcritos por Pol III; por ejemplo, en el cromosoma 23 (Fig. 13, Tabla 9) hay 10 genes de RNAt, un gen de RNAr 5S y dos genes de RNAsn: U1 y U3. Otro ejemplo es el locus IV en el cromosoma 36 (Chr36-IV), que posee 4 genes RNAt y el gen RNAsn U5. Cabe mencionar que únicamente 8 *loci* contienen sólo un gen de RNAt (Chr03, Chr06, Chr07, Chr16, Chr30, Chr24-I, Chr24-III y Chr36-III).

En general, las regiones intergénicas que separan los genes de Pol III son cortas, con un tamaño promedio de 202 bases (Fig. 13). La región intergénica más pequeña, de sólo 35 bases, es la que separa los genes de RNAt-Met de elongación y RNAt-Leu en el *cluster* Chr34-I, y la más grande (de 5406 bases) es la localizada entre RNAt-Gly y RNAt-Ala, en el *cluster* Chr11-I.

En *T. brucei*, los 66 genes de RNAt están repartidos en 26 *loci*, en 8 cromosomas diferentes (Fig. 14, Tabla 9). De manera similar a *L. major*, en *T. brucei* el número de genes de RNAt por *cluster* es de 2-10 genes, y 11 de los 66 genes se presentan como un RNAt único en su locus correspondiente: chr04-I, chr05-I, chr05-II, chr07-I, chr07-III, chr07-IV, chr08-IV, chr09-I, chr09-II, chr09-III y chr11-I. Como ocurre en *L. major*, las regiones intergénicas entre los genes de Pol III son cortas, con 327 bases en promedio y un rango de 43 a 3172 bases (en los *clusters* chr04-IV y chr07-II, respectivamente, Fig. 14).

Con respecto a *T. cruzi*, no fue posible reportar la posición relativa de los genes de RNAt y otros genes de Pol III, ya que su genoma aún no ha sido ensamblado en cromosomas y se presentan sólo pequeños fragmentos en las bases de datos.

Tabla 9: Distribución cromosomal de genes de RNAt en *L. major* y *T. brucei* (Cr, cromosoma).

<i>Leishmania major</i>					
Cr	Cluster	Genes RNAt	Otros genes transcritos por Pol III	Total de genes	Localización del cluster en relación con genes transcritos por Pol II
3	I	1	-	1	Convergente
5	I	1	1 7SL, 1 5S	3	Convergente
6	I	1	-	1	Dentro de DGC
7	I	1	-	1	Dentro de DGC
9	I	2	1 5S	3	Convergente
	II	3	1 5S	4	Divergente
	III	4	1 5S	5	Convergente
10	I	3	-	3	Divergente
11	I	5	2 5S	7	Dentro de DGC
	II	2	1 5S	3	Dentro de DGC
15	I	2	1 5S	3	Convergente
16	I	1	-	1	Convergente
17	I	3	-	3	Dentro de DGC
21	I	1	1 5S	2	Convergente
	II	3	1 5S	4	Convergente
23	I	10	1 5S, 1 U1, 1 U3	13	Convergente
24	I	1	-	1	Dentro de DGC
	II	4	1 U6	5	Convergente
	III	1	-	1	Dentro de DGC
29	I	2	-	2	Dentro de DGC
30	I	1	-	1	Convergente
31	I	1	1 U2	2	Dentro de DGC
	II	2	-	2	Dentro de DGC
	III	2	-	2	Dentro de DGC
33	I	3	-	3	Dentro de DGC
34	I	5	-	5	Convergente
	II	5	-	5	Dentro de DGC
36	I	4	-	4	Convergente
	II	4	1 U4	5	Convergente
	III	1	-	1	Dentro de DGC
	IV	4	1 U5	5	Dentro de DGC
Total genes tRNA		83			

<i>Trypanosoma brucei</i>					
Cr	Cluster	Genes RNAt	Otros genes transcritos por Pol III	Total de genes	Localización del cluster en relación con genes transcritos por Pol II
3	I	3	-	3	Dentro de DGC
4	I	1	-	1	divergente
	II	2	1 U6	3	Dentro de DGC
	III	3	-	3	convergente
	IV	4	-	4	Dentro de DGC
5	V	1	1 sRNA76	2	Dentro de DGC
	I	1	-	1	Dentro de DGC
7	II	1	-	1	convergente
	I	1	-	1	Dentro de DGC
	II	6	-	6	Dentro de DGC
	III	1	-	1	Dentro de DGC
8	IV	1	-	1	Dentro de DGC
	I	2	-	2	Dentro de DGC
	II	10	1 U1, 1 7SL, 1 U3	13	convergente
	III	2	-	2	convergente
	IV	1	-	1	convergente
9	V	5	-	5	convergente
	I	1	-	1	Dentro de DGC
	II	1	-	1	Dentro de DGC
	III	1	-	1	Dentro de DGC
	I	1	1 U4	2	Dentro de DGC
10	II	5	1 U5	6	convergente
	III	5	-	5	convergente
	I	1	-	1	Dentro de DGC
11	II	3	-	3	Dentro de DGC
	III	3	-	3	convergente
	Total genes tRNA		66		

5.5. Relación espacial entre los genes de Pol II y Pol III en tripanosomátidos.

L. major posee, en el cromosoma 3, dos DGCs convergentes separados por un gen de RNAt-Lys. El análisis transcripcional llevado a cabo en dicho cromosoma indicó que la transcripción de los dos DGCs, llevada a cabo por Pol II, termina en el área de convergencia (donde hay un cambio de la hebra codificadora) dentro de la región del gen de RNAt (Martínez-Calvillo *et al*, 2004). Al analizar los mapas de cromosomas reportados en GeneDB encontramos 39 regiones de cambio de hebra convergentes en el genoma de *L. major*. Catorce de ellas (37%) contienen al menos un gen de RNAt (Fig. 13). Esto sugiere que en otros DGCs convergentes de genes codificantes para proteínas (además de los localizados en el cromosoma 3), los RNAt podrían usarse como una señal para el término de la transcripción en tripanosomátidos.

De 31 *loci* que contienen genes de RNAt en *L. major*, 14 (45%) se localizan en zonas convergentes de cambio de hebra codificante (Fig. 13, Chr 03) y 2 (6%) en áreas divergentes (Fig. 13, Chr10). Los restantes 15 *loci* de RNAt (48%) están situadas dentro de DGCs (Fig. 13, Chr07). Algo similar se encontró en *T. brucei*: 36% de los *clusters* de RNAt se localizan entre dos DGCs convergentes (Fig. 14, Chr04-III), 4% entre zonas divergentes (Fig. 14, Chr04-I) y 60% de ellos dentro de DGCs, en la misma hebra codificadora (Fig. 14, Chr08-I) (Tablas 8 y 9).

Tabla 10. Localización de los *clusters* de RNAt en relación a las unidades de genes de Pol II.

	<i>L. major</i> ¹	<i>T. brucei</i> ²
Entre unidades Convergentes	45%	36%
Entre unidades Divergentes	6%	4%
Dentro de unidades	48%	60%

¹ 31 *clusters* de RNAt (Fig. 13)

² 27 *clusters* de RNAt (Fig. 14)

5.6. Sintenia entre genes de Pol III

Se ha reportado que los genomas de los tripanosomátidos son altamente sinténicos entre sí, es decir, que muestran una elevada conservación en el orden de genes homólogos entre fragmentos de cromosomas. La conservación en el orden de los genes ha sido observada en DGCs, que son transcritos por Pol II. Quisimos determinar si los genes de Pol III eran sinténicos entre los tripanosomátidos, por lo que se analizaron dichas secuencias. Encontramos que la mayoría de los *clusters* de RNAt no muestran sintenia entre sí. Sin embargo, varios *loci* muestran un elevado nivel de conservación (Fig. 19); El caso más representativo es un *cluster* de 13 genes de Pol III, Chr23 de *L. major*, el *cluster* Chr08-II en *T. brucei* y un *cluster* localizado en el contig Tc6288 en *T. cruzi* (Fig. 19A). El orden de los genes en este *cluster* es idéntico entre *T. cruzi* y *T. brucei*, con la diferencia de que el gen RNAsn U1, el RNA 7SL y el RNAt-Leu se localizan en hebras diferentes. La mayoría de estos genes se encuentran en el *cluster* de *L. major*, pero su orden no es idéntico; además, en *L. major* se encuentra un gen de RNAr 5S que reemplaza a un gen de RNA 7SL, y un gen de RNAt-Trp que reemplaza al gen de RNAt-Lys (Fig. 19A).

El siguiente *cluster* que presenta sintenia es el Chr24-II en *L. major*, Chr08-V en *T. brucei* y el contig Tc6223 de *T. cruzi* (Fig. 19B), en los que se encuentran genes de Ile, Leu y Gln sinténicos entre las tres especies. En *L. major* y *T. brucei* se encuentra una segunda copia de RNAt-Ile y en *T. brucei* una segunda de RNAt-Gln, mientras que *L. major* posee además un gen del RNAsn U6. Otros *clusters* sinténicos de genes de RNAt son los *clusters* Chr33-I de *L. major*, Chr10-II de *T. brucei* y el contig Tc8001 de *T. cruzi* (Fig. 19C) y Chr34-I de *L. major*, Chr04-III de *T. brucei* y contig Tc4886 de *T. cruzi* (Fig. 19D). Varios genes de proteínas que flanquean estos *clusters* de RNAt son también sinténicos entre los tripanosomátidos.

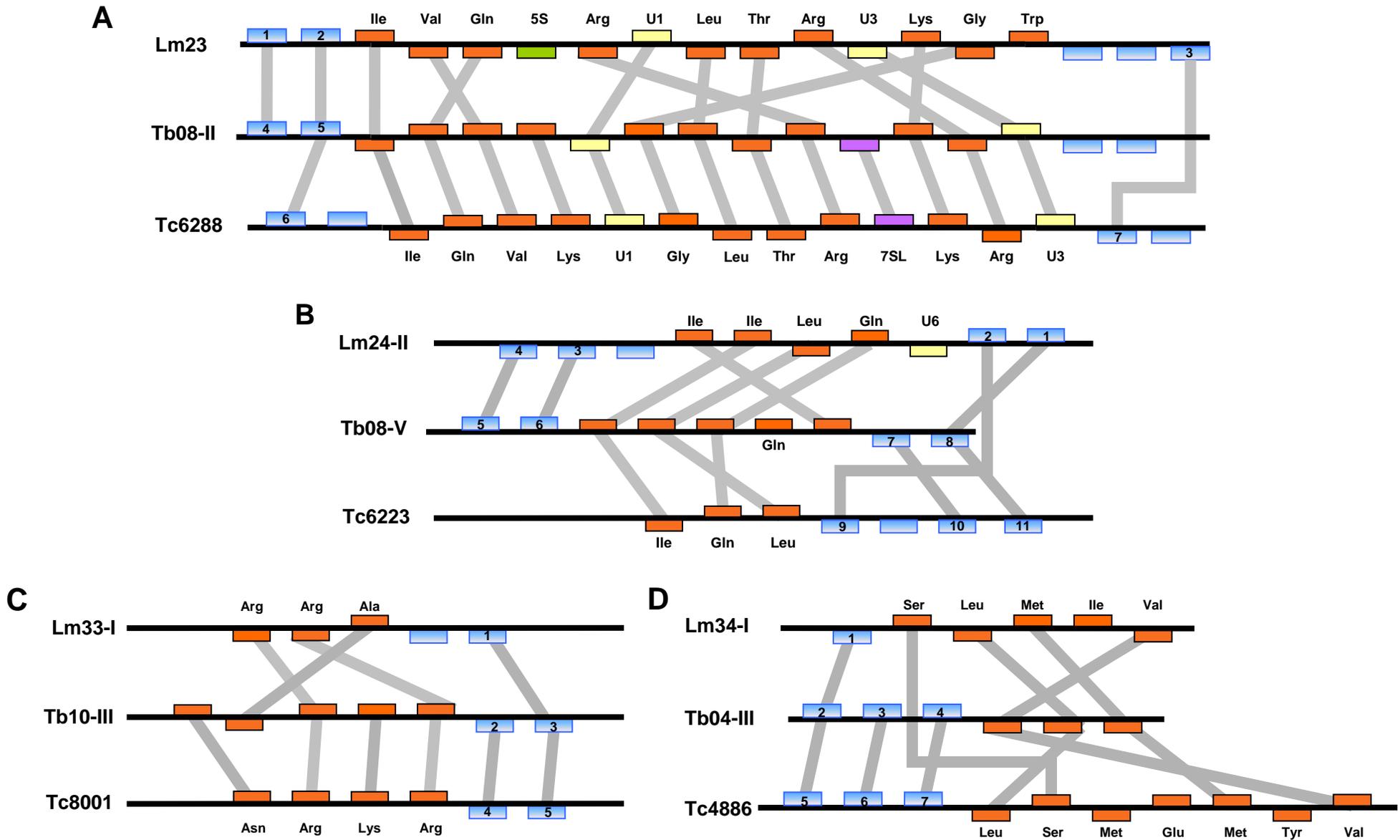


Figura 19. Sintenia de genes transcritos por Pol III en Tritryps. Se muestra los *clusters* que presentan cierto grado de sintenia entre Tritryps. En *L. major* corresponden al *loci* Chr23 (A), Chr24-II (B), Chr33-I (C) y Chr34-I (D). El orden de los genes en los *clusters* Chr24-II y Chr34-I se invirtió (con respecto a la Fig. 13). Los genes homólogos están unidos por una barra gris.

5.7. Intrones en los genes de RNAt en tripanosomátidos.

El análisis de las secuencias de los genes de RNAt en tripanosomátidos reveló que los genes de RNAt de tirosina, anticodón GTA, son los únicos que poseen un intrón dentro de su estructura. El intron en dicho gen ya había sido reportado en el caso de *T. brucei* (Schneider *et al*, 1993).

El intrón posee 11 bases en *L. major* y *T. brucei*, mientras que el de *T. cruzi* tiene 13 pb y, de manera similar a lo encontrado en otros organismos (Marck *et al*, 2002), comienza en el nucleótido 38 del RNAt (Fig. 20). Cabe mencionar que los genes de RNAt-Tyr entre tripanosomátidos exhiben una similitud muy alta al comparar las secuencias completas de los genes entre las tres especies (el porcentaje de identidad va del 94-97%); sin embargo, las diferencias entre las secuencias se limitan exclusivamente a la región del intron, pues el resto de las secuencias de los genes de RNAt-Tyr son idénticas (Fig. 20).

		intron	
LmjF36.TRNATYR.01	CCTTCTGTAGCTCAATTGGTAGAGCATGTGACTGTAGAGTTC--GCGGGT		48
LmjF36.TRNATYR.02	CCTTCTGTAGCTCAATTGGTAGAGCATGTGACTGTAGAGTTC--GCGGGT		48
LmjF34.TRNATYR.01	CCTTCTGTAGCTCAATTGGTAGAGCATGTGACTGTAGAGTTC--GCAGGT		48
Tb927.4.1219	CCTTCTGTAGCTCAATTGGTAGAGCATGTGACTGTAGAGTAT--GCGGGT		48
Tc00.1047053504055.93	CCTTCTGTAGCTCAATTGGTAGAGCATGTGACTGTAGAAATTCGGCCGGTA		50
Tc00.1047053506435.354	CCTTCTGTAGCTCAATTGGTAGAGCATGTGACTGTAGAACTTCGGCCGGTA		50
	*****	** *	
LmjF36.TRNATYR.01	ATCACAGGGTCGCTGGTTCGATTCCGGCCGGAAGGA		84
LmjF36.TRNATYR.02	ATCACAGGGTCGCTGGTTCGATTCCGGCCGGAAGGA		84
LmjF34.TRNATYR.01	ATCACAGGGTCGCTGGTTCGATTCCGGCCGGAAGGA		84
Tb927.4.1219	ATCACAGGGTCGCTGGTTCGATTCCGGCCGGAAGGA		84
Tc00.1047053504055.93	ATCACAGGGTCGCTGGTTCGATTCCGGCCGGAAGGA		86
Tc00.1047053506435.354	ATCACAGGGTCGCTGGTTCGATTCCGGCCGGAAGGA		86

Figura 20. Alineamiento de los genes de RNAt de tirosina en los tres tripanosomátidos. *L. major* posee tres genes de Tyr, *T. brucei* uno y *T. cruzi* dos. * significa que las secuencias poseen el mismo nt en esa misma posición.

5.8. Secuencias consenso de los elementos promotores

El promotor de los genes de RNAt consiste en dos elementos, las cajas A y B, que se localizan dentro de la secuencia del gen (Fig. 2). Se identificaron los probables elementos de control interno en las secuencias de todos los genes de RNAt de los tres tripanosomátidos y se obtuvieron secuencias consenso para ambos elementos promotores, tomando como referencia la secuencia de las cajas A y B de *S. cerevisiae*.

Al realizar el análisis de las secuencias se identificaron en los tripanosomátidos las dos clases de RNAt (clases I y II) reportadas para otros organismos (Brennan y M. Sundaralingam, 1976) (Figs. 5C, 5D y 18). La clase I posee un asa variable corta de 4 ó 5 nt, mientras que la clase II posee un brazo variable largo. Entre los tripanosomátidos, 226 genes pertenecen a la clase I y 43 genes pertenecen a la clase II (todos los RNAt de Leu y Ser, excluyendo los de Sec, cuyas particularidades se discuten más adelante). Además de las diferencias en el tamaño del brazo variable, encontramos que las dos clases difieren entre sí en las secuencias de los elementos promotores; por esto se reportan secuencias consenso para cada una de las clases de RNAt de tripanosomátidos (Fig. 21).

Caja A												
	1	2	3	4	5	6	7	8	9	*	10	11
Tritryp clase I	T	A/G	G	C/T	N	C/T	A	G/A	T/C	C/T	G	G
Tritryp clase II	T	G/A	G/C	C/T	C	G/A	A	G	T	T	G	G
<i>S. cerevisiae</i>	T	G/A	G	C/T	N	N	A	G/A	H	H	G	G

Caja B											
	1	2	3	4	5	6	7	8	9	10	11
Tritryp clase I	G/A	G	T	T	C	G/A	A	N	H	C	C/T
Tritryp clase II	G	G	T	T	C	G/A	A	A	C/T	C	C
<i>S. cerevisiae</i>	N	G	T	T	C	G/A	A	N	C/T	C	N

Figura 21. Secuencias consenso de las cajas A y B de los genes de RNAt en Tritryps. Las secuencias consenso de las clases I y II se comparan contra la reportada para levadura. En la caja A, algunos RNAt poseen una base extra entre las bases 9 y 10 (marcada con un asterisco). Las posiciones conservadas se muestran en negritas. En caso de que dos nucleótidos sean posibles, se escribe primero el que ocurre con mayor frecuencia. H= A, C, T.

Alrededor del 20% de los genes de la clase I poseen el nucleótido extra entre las posiciones 9 y 10, que corresponde a los mismos genes en las tres especies (entre paréntesis el anticodón): Met- (CAT), Asn (GTT), Ile (AAT), Ile (TAT), Lys (CTT), Phe (GAG), Tyr (GTA) y sólo en *L. major*, Lys (TTT). Los genes de clase II que poseen un nucleótido en esta posición son aquellos de Leu (TAA) en las tres especies. Otras excepciones fueron encontradas en la caja A de algunos RNAt clase I (ver Tabla 10A), en la caja A de varios RNAt clase II (Tabla 10B) y en la caja B de algunos RNAt de clase I (Tabla 10C).

Tabla 11. Excepciones a la secuencia consenso de las cajas promotoras.

Tabla 11-A. Caja A de genes RNAt Clase I

Posición	Organismo ^b	Genes	Descripción
2	Lm Tb Tc	LmjF09.VAL.01 (CAC) LmjF09.VAL.02 (CAC) Tb08_tRNA_Val_1 (Tb927.8.2853) (TAC) Tc00.1047053506459.249 (Val-CAC)	Poseen una C en posición 2 en lugar de A/G
3	Lm Tb Tc	LmjF23.VAL.01(TAC) Tb08_tRNA_Val_1 (Tb927.8.2853) (TAC) Tc00.1047053504427.233 (VAL-TAC) Tc00.1047053508043.13 (Val-TAC)	Poseen A en lugar de G

Tabla 11-A. Continuación

<p>Entre 8 y 9^a</p>	<p>Lm Tb Tc</p>	<p>LmjF17.TRNAALA.01 (AGC) LmjF31.TRNAALA.01 (AGC) Tb07_tRNA_Ala_1 (Tb927.7.6821) (AGC) Tb07_tRNA_Ala_2 (Tb927.7.6826) (AGC) Tc00.1047053510057.40 (ALA-AGC) Tc00.1047053508909.130 (Ala, AGC)</p>	<p>Poseen una A entre estas posiciones. Característica exclusiva de los genes con este isotipo</p>
<p>Entre 8 y 9^a</p>	<p>Lm Tb Tc</p>	<p>LmjF21.TRNAVAL.01 (AAC) LmjF34.TRNAVAL.01 (AAC) Tb04_tRNA_Val_1 (Tb927.4.3193) (AAC) Tb04_tRNA_Val_2 (Tb927.4.3298) (AAC) Tc00.1047053506435.363 (VAL-AAC) Tc00.1047053504055.95 (Val-AAC)</p>	<p>Poseen una G entre estas posiciones. Característica exclusiva de los genes con este isotipo.</p>
<p>* (entre las posiciones 9 y 10)</p>	<p>Lm (22% de los genes RNAt) Tb (20%) Tc (22%)</p>	<p>Genes que contienen C esta posición ASN – GTT Tb10_tRNA_Asn_1, Tb08_tRNA_Asn_1 (Tb927.8.2663), Tc00.1047053503865.4 , Tc00.1047053475029.50, Tc00.1047053508693.110, Tc00.1047053509105.112 LmjF10.TRNAASN.01, LmjF34.TRNAASN.01, LmjF34.TRNAASN.02</p> <p>ILE- AAT Tb08_tRNA_Ile_3 (Tb927.8.6572), Tb08_tRNA_Ile_2 (Tb927.8.6562), Tc00.1047053511909.3, Tc00.1047053508777.205, Tc00.1047053511483.70, Tc00.1047053510721.19, LmjF24.TRNAILE.01, LmjF24.TRNAILE.02, LmjF34.TRNAILE.01</p> <p>LYS-CTT Tb08_tRNA_Lys_1 (Tb927.8.2854), Tb08_tRNA_Lys_2 (Tb927.8.2862), Tb07_tRNA_Lys_1 (Tb927.7.6845). Tc00.1047053504427.234, Tc00.1047053504427.242, Tc00.1047053508043.22, Tc00.1047053508043.14 LmjF03.TRNALYS.01, LmjF10.TRNALYS.01, LmjF21.TRNALYS.02</p> <p>Met-e Tb04_tRNA_Met_1 (Tb927.4.3199), Tb04_tRNA_Met_2 (Tb927.4.3294) Tc00.1047053506435.345, Tc00.1047053504055.91, Tc00.1047053506435.327, Tc00.1047053504055.87, LmjF11.TRNAMET.01, LmjF34.TRNAMET.01</p>	<p>Genes que poseen C en la posición opcional entre la 9 y la 10 (mencionado en el texto)</p>
		<p>PHE-GAA Tb11_tRNA_Phe_1, Tb11_tRNA_Phe_2, Tc00.1047053506599.9, Tc00.1047053508873.402, Tc00.1047053506459.243, Tc00.1047053506603.60, LmjF09.TRNAPHE.01, LmjF31.TRNAPHE.01</p> <p>TYR-GTA Tb04_tRNA_Tyr_1 (Tb927.4.1219), Tc00.1047053504055.93, Tc00.1047053506435.354, LmjF34.TRNATYR.01, LmjF36.TRNATYR.01, LmjF36.TRNATYR.02</p> <p>ILE- TAT Tb08_tRNA_Ile_1 (Tb927.8.2851), Tc00.1047053504427.231, Tc00.1047053508043.11, LmjF23.TRNAILE.01</p> <p>LYS- TTT (Solo en Lm):LmjF23.TRNALYS.01</p>	<p>Genes que poseen T en la posición opcional entre la 9 y 10</p>

^b Lm, *L. major*, Tb, *T. brucei* y Tc, *T. cruzi*.

Tabla 11-B. Caja A de genes RNAt Clase II

Posición	Organismo ^b	Genes	Descripción
2	<i>L. major</i> <i>T. brucei</i> <i>T. cruzi</i>	LmjF24.TRNALEU.01 (TAA) Tb08_tRNA_Leu_2 (Tb927.8.6564) (TAA) Tc00.1047053510721.13 (Leu TAA) Tc00.1047053511909.9 (Leu TAA)	Poseen T en lugar de G/A
* (entre 9 y 10)	<i>L. major</i> <i>T. brucei</i> <i>T. cruzi</i>	LmjF24.TRNALEU.01 (TAA) Tb08_tRNA_Leu_2 (Tb927.8.6564) TAA Tc00.1047053510721.13 (Leu TAA) Tc00.1047053511909.9 (Leu) TAA	Poseen T en esta posición opcional

^b Lm, *L. major*, Tb, *T. brucei* y Tc, *T. cruzi*.

Tabla 11-C. Caja B de genes Clase I

Posición	Organismo	Gene	Descripción
4	<i>L. major</i>	LmjF09.TRNAHIS.01 (His-GTG) LmjF09.TRNAHIS.02 (His-GTG)	Poseen C en esta posición en lugar de T
5	<i>T. brucei</i>	Tb07_tRNA_Ala_3 (Tb927.7.7175) (Ala-CGC)	Único gen que tiene T en esta posición en lugar de C
10	<i>T. cruzi</i>	Tc00.1047053511241.10 (His-GTG)	Único que tiene A en esta posición en lugar de C

Al analizar los genes de RNAt-Sec de los tripanosomátidos, se encontró que la caja A contiene una A adicional entre las bases 2 y 3, cuando se compara su secuencia con la secuencia consenso (Fig. 22A). Esta inserción ha sido previamente reportada en otros organismos (Lee *et al*, 1989). En el caso de la caja B, los genes de RNAt-Sec de tripanosomátidos presenta dos cambios respecto a la secuencia consenso de los RNAt convencionales: una C en posición 1 (en lugar de G) y G en posición 11 (en lugar de C/T) (Fig. 22B). Esto difiere de lo observado en otros eucariontes, en los que la caja B de los RNAt-Sec es idéntica a la secuencia consenso del resto de los RNAt.

A

<i>H.sapiens</i>	T G <u>A</u> T C C T C A G T G G
<i>X.laevis</i>	T G <u>A</u> C C C T C A G T G G
<i>D.melanogaster</i>	T G <u>A</u> A C T T C G G T G G
<i>C.elegans</i>	T G <u>A</u> A C C A T G G C G G
<i>P.falciparum</i>	T G <u>A</u> G T T A G C A T G G
Trityps Sec	T G <u>A</u> G C T C A G C T G G
	* * * * *
Consenso Trityps (RNAt clásicos)	T G G C N C A G C T G G

B

<i>H.sapiens</i>	G G T T C A A T T C C
<i>X.laevis</i>	G G T T C A A T T C C
<i>D.melanogaster</i>	G G T T C G A T T C C
<i>C.elegans</i>	G G T T C G A C T C C
<i>P.falciparum</i>	G G T T C G A T T C C
Trityps Sec	C G T T C G A T T C G
	* * * * * * *
Consenso Trityps (RNAt clásicos)	G G T T C G A N T C CT

Figura 22. Comparación de las cajas promotoras en los genes de RNAt-Sec. A) Comparación de las cajas A entre eucariontes. A en la tercera posición (en negritas y subrayada) es una inserción en la caja en estos genes y que no se encuentra en las cajas de los RNAt convencionales. B) Comparación de la caja B entre especies. Los nucleótidos conservados se marcan en ambas cajas con un asterisco, se incluye en gris la secuencia consenso para las cajas A y B de los genes de tRNA clásicos (clase I) en Trityps.

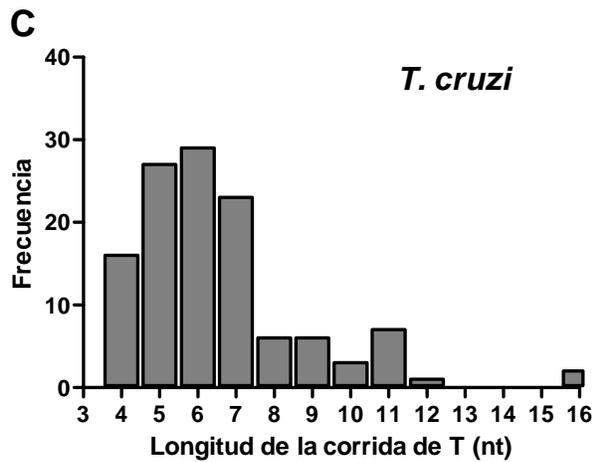
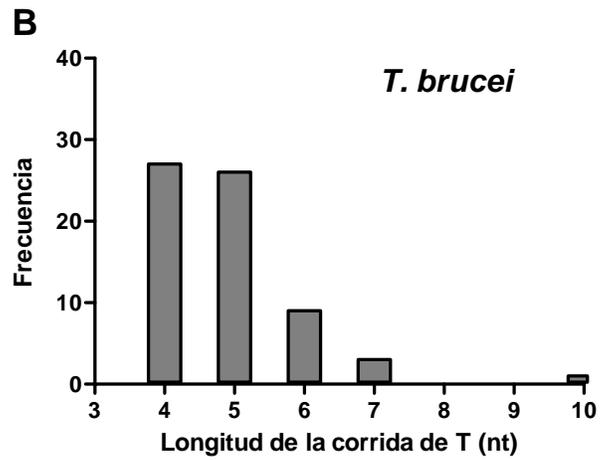
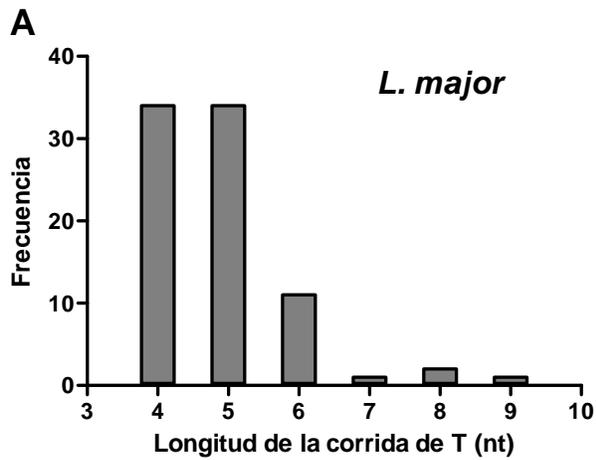
5.9. Señales de término de la transcripción en genes de RNAt.

Una secuencia de varios residuos de T que se localiza río abajo del extremo 3' de los genes de RNAt actúa como señal para terminar la transcripción por Pol III (Fig. 2). El *cluster* de Ts se encuentra generalmente dentro de las primeras 40 bases después del gen. Para saber más sobre las secuencias de terminación de la transcripción en tripanosomátidos, analizamos las secuencias 3' de todos los genes de RNAt y realizamos un análisis estadístico. Se encontró una secuencia de Ts de longitud variable en cada uno de los genes (Apéndice II); la distancia entre el extremo del gen y la secuencia de Ts va de cero a siete bases.

En *L. major*, el promedio de la longitud de la señal de término es de 4.87 bases y comprende de 4 a 9 residuos (Fig. 23A). Algo similar se encontró en *T. brucei*, en donde la secuencia de Ts es de

4.9 bases en promedio, con un rango de 4 a 10 bases (Fig. 23B). El patrón de las señales de término en *T. cruzi* es diferente de las primeras dos especies ya que las secuencias son más largas, con una media de 6.55 residuos y un máximo de 16 (Fig. 23C).

La presencia de una segunda secuencia de Ts (“back up”) que sirva como refuerzo de la señal de terminación es una característica común de los RNAt en eucariontes (Braglia et al, 2005). Esta segunda secuencia se encuentra generalmente 30 pb río abajo de la primera. El análisis de las secuencias en tripanosomátidos, reveló que *T. cruzi* posee 58 genes de RNAt (48%) que presentan la secuencia “back up”, cuya longitud va de 4-15 bases. Algo similar se ha observado en otros eucariontes como *S. cerevisiae* y *S. pombe*, cuyos porcentajes de presencia de back-up son 44 y 53% respectivamente (Fig. 24) (Braglia et al, 2005). Las otras dos especies mostraron un comportamiento diferente, pues *L. major* presentó sólo 13 genes de RNAt con la segunda secuencia (16%) y *T. brucei* 18 genes (27%), conformadas por 4-10 bases (Fig. 24) (Apéndice II).



D

	Lm	Tb	Tc
Promedio	4.87	4.89	6.57
Desv. Est.	1.0	1.06	2.27
Min.	4	4	4
Max.	9	10	16
N	83	66	120

Figura 23. Distribución de la longitud de la señal de término de los genes de RNAt en Trityps. Se muestra el histograma del tamaño de la secuencia de Ts en cada gen de RNAt de *L. major* (panel A), *T. brucei* (panel B) y *T. cruzi* (panel C). Los datos estadísticos se muestran en la tabla D (Desv. Est = desviación estándar, Min, número mínimo de residuos en la secuencia, Max., máximo número de residuos, N, número de genes).

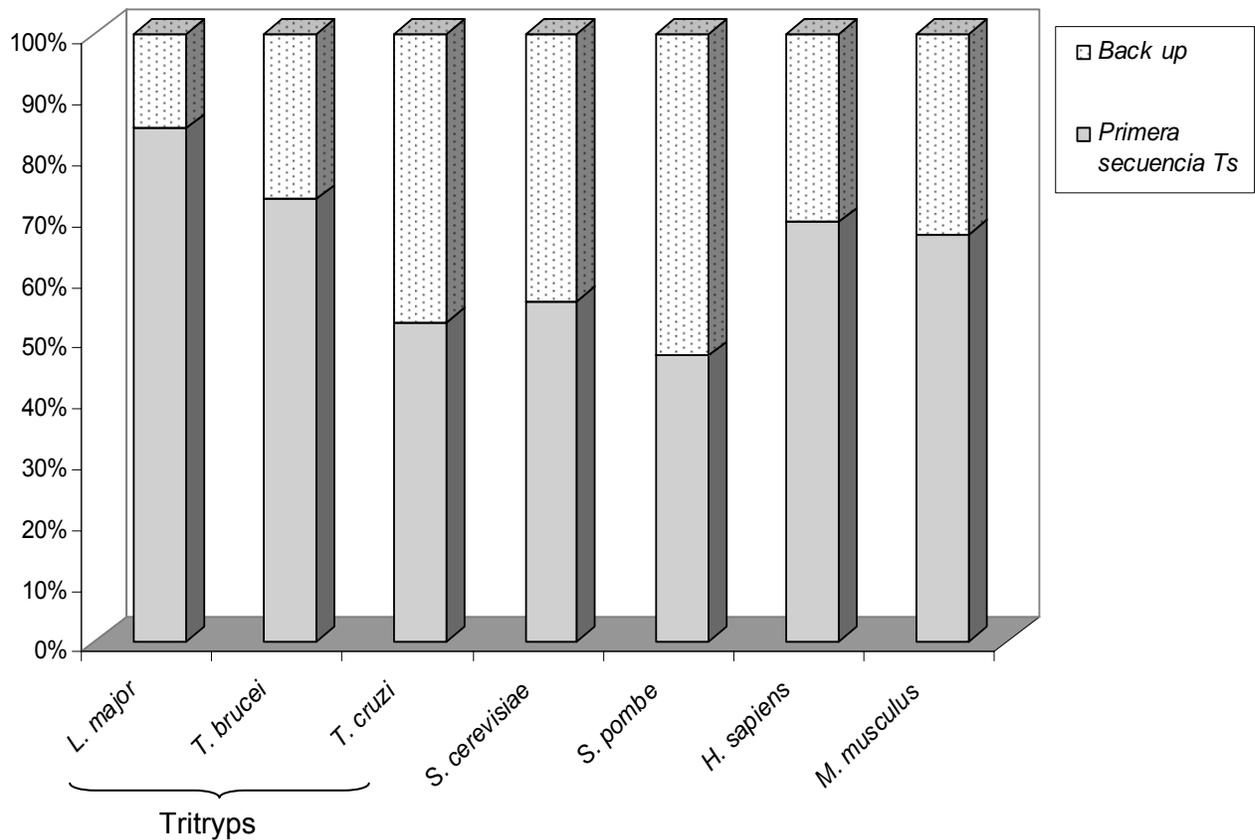


Figura 24. Porcentaje de genes de RNAt que presentan una segunda señal de término de la transcripción en Trityps y otros eucariontes. La barra inferior corresponde al porcentaje de genes que sólo tienen una secuencia de Ts y la barra superior punteada indica el porcentaje de genes de tRNA que presentan una segunda secuencia de Ts (*back up*). Nótese que *T. cruzi* presenta un porcentaje similar al de *S. cerevisiae* y *S. pombe*.

5.10. RNAt isodecodificadores.

Al analizar secuencias de RNAt isoaceptores (aquellos que acarrean el mismo aminoácido pero con diferente anticodón) se observó la presencia de RNAt isodecodificadores, es decir, RNAt que codifican para el mismo anticodón, pero poseen diferencias en el resto de la secuencia (Goodenbour y Pan, 2006). Quisimos saber si existían RNAt isodecodificadores en los Trityps y cuales eran esos genes. Dada una clase específica de RNAt isoaceptor, la secuencia que posee el mayor número de copias génicas se designa arbitrariamente como el “miembro mayoritario”. El número de genes isodecodificadores para una clase de isoaceptor es el número de secuencias de

RNA^t diferentes, excluyendo al “miembro mayoritario”. Así, se encontró 1 isodecodificador para *L. major* y 1 para *T. brucei* (Fig. 25). Ya que solo existen dos copias del isoceptor correspondiente en cada caso, designamos arbitrariamente a los genes LmtRNA-Ser.01 (LmjF17.TRNASER.01) y TbtRNA-Ala.01 (Tb07_tRNA_Ala_3) como los isodecodificadores. La identidad de secuencia entre el isodecodificador y el “miembro mayoritario” es de 98% en *L. major* y 97% en *T. brucei* (Fig. 25 y Tabla 11).

En *T. cruzi* se identificaron 6 genes isodecodificadores, los cuales se enumeran en la Tabla 11. El porcentaje de identidad entre el “miembro mayoritario” y el isodecodificador es muy alto, entre el 95-98% (Tabla 11). En *L. major*, las diferencias entre las secuencias se localizan cerca de la caja A, mientras que en *T. brucei*, una de las diferencias se localiza dentro de la caja B en una posición conservada (Fig. 25). En el caso de *T. cruzi*, en cuatro isodecodificadores las diferencias en las secuencias caen dentro de la caja B, en nucleótidos no conservados: RNA^t-Ala (TGC), RNA^t-Glu, RNA^t-Arg, RNA^t-His (Fig. 25).

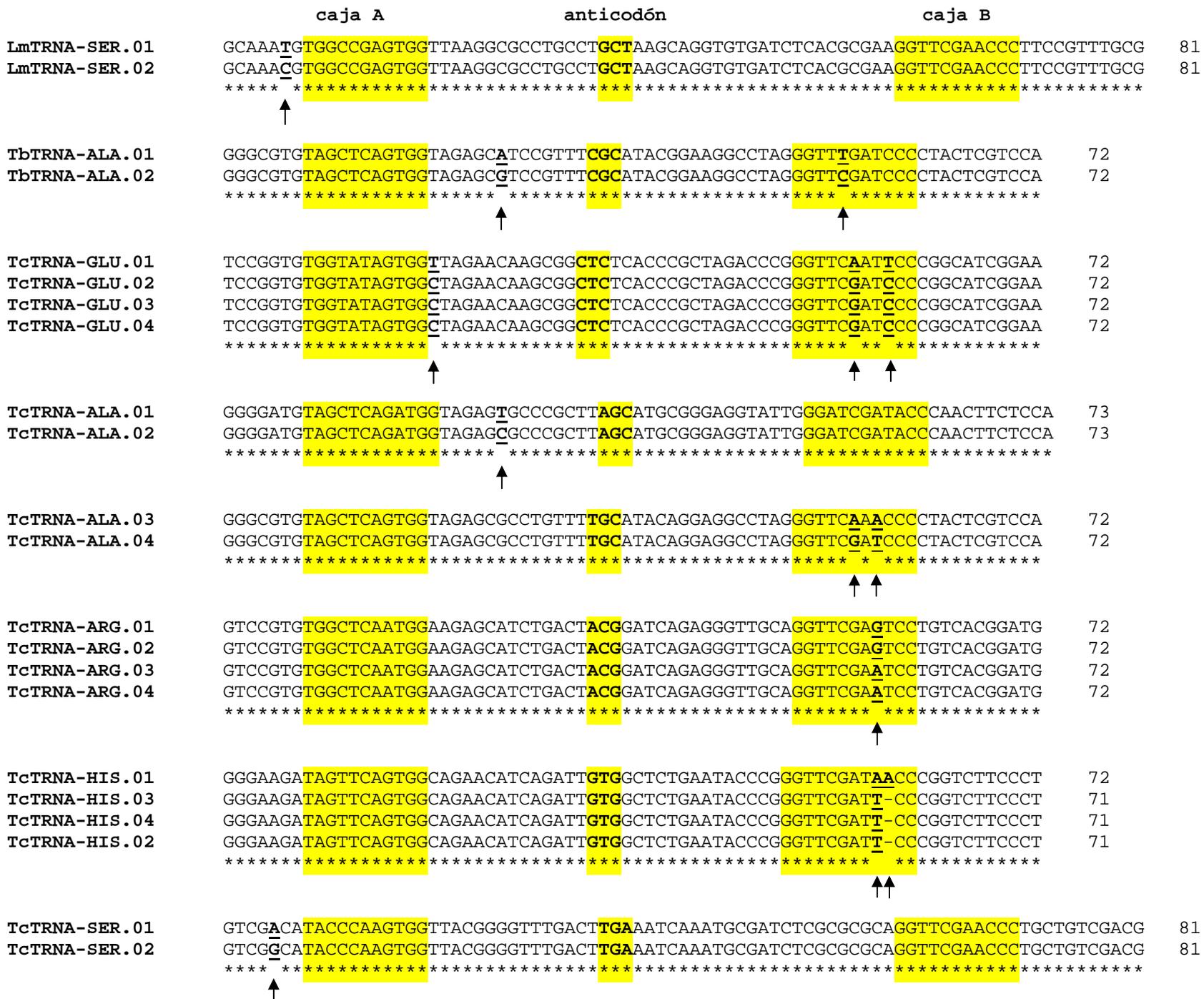


Figura 25. RNAt isodecodificadores en Trityps. Pie de figura página siguiente

Figura 25. RNAt isodecodificadores en Trityps. Se encontró un isodecodificador en *L. major* (Ser-GCT), uno en *T. brucei* (Ala-CGC) y seis en *T. cruzi* (RNAt-Glu, -Ala, -Arg, -His y -Ser). Resaltado en amarillo se indican las cajas A y B y el anticodón. Las bases que varían se muestran en negritas, subrayadas y marcadas con una flecha. Los genes son los siguientes, nombre de GeneDB en paréntesis: LmjF17.TRNASER.01 (LmjF17.TRNASER.01), LmjF21.TRNASER.01 (LmjF21.TRNASER.01), TbTRNA-ALA.01 (Tb07_tRNA_Ala_3 también Tb927.7.7175), TbTRNA-ALA.02 (Tb11_tRNA_Ala_1), TcTRNA-GLU.01 (Tc00.1047053506435.336), TcTRNA-GLU.02 (Tc00.1047053504055.89), TcTRNA-GLU.03 (Tc00.1047053508999.180), TcTRNA-GLU.04 (Tc00.1047053510959.8), TcTRNA-ALA.01 (Tc00.1047053510057.40), TcTRNA-ALA.02 (Tc00.1047053508909.130), TcTRNA-ALA.03 (Tc00.1047053475029.40), TcTRNA-ALA.04 (no anotado, contig 8001) TcTRNA-ARG.01 (Tc00.1047053504427.243), TcTRNA-ARG.02 (Tc00.1047053506619.59), TcTRNA-ARG.03 (Tc00.1047053508043.23), TcTRNA-ARG.04 (Tc00.1047053511191.29), TcTRNA-HIS.01 (Tc00.1047053511241.10), TcTRNA-HIS.03 (Tc00.1047053508087.5), TcTRNA-HIS.04 (Tc00.1047053508861.10), TcTRNA-HIS.02 (Tc00.1047053506663.10), TcTRNA-SER.01 (Tc00.1047053510057.50), TcTRNA-SER.02 (Tc00.1047053508909.120)

Tabla 12. Genes RNAt isodecodificadores

	<i>L. major</i>	<i>T. brucei</i>	<i>T. cruzi</i>
Miembro mayoritario	LmjF21.TRNASER.01 (Ser-GCT)	Tb11_tRNA_Ala_1 (Ala-CGC)	ALA (AGC) Tc00.1047053508909.130 ALA (TGC) Tc00.1047053475029.40 ARG (ACG) Tc00.1047053506619.59 Tc00.1047053508043.23 Tc00.1047053511191.29 HIS (GTG) Tc00.1047053508087.5 Tc00.1047053508861.10 Tc00.1047053506663.10 SER (TGA) Tc00.1047053508909.120 GLU (CTC) Tc00.1047053504055.89 Tc00.1047053508999.180 Tc00.1047053510959.8
Isodecodificador	LmjF17.TRNASER.01 (Ser-GCT – 98% ¹)	Tb07_tRNA_Ala_3 (Tb927.7.7175) Ala- CGC – 97% ¹	ALA (AGC) – 98% ¹ Tc00.1047053510057.40 ALA (TGC) – 97% ¹ Contig 8001 (28744-28815) ARG (ACG) – 98% ¹ Tc00.1047053504427.243 HIS (GTG) – 98% ¹ Tc00.1047053511241.10 SER (TGA) – 98% ¹ Tc00.1047053510057.50 TcTRNA-GLU – 95% ¹ Tc00.1047053506435.336

¹ Al lado de cada isoceptor se menciona el porcentaje de identidad de secuencia que existen entre el miembro mayoritario y el isodecodificador.

5.11. Uso de codones en tripanosomátidos

Como se mencionó anteriormente, en muchos organismos existe una correlación entre el número de copias génicas de RNAt y el uso de codones (Percudani *et al*, 1997; Rocha, 2004). Aparentemente, la selección de codones sinónimos conduce a una coadaptación entre el uso de codones y el contenido de RNAt para optimizar la síntesis de proteínas. Para conocer el escenario en tripanosomátidos, se realizó un análisis de correlación de Spearman entre el uso de codones y el número de RNAt por isoceptor. Se incluyeron todos los genes codificadores de proteínas de *L. major* (8272 genes, 5.2×10^6 codones), 5140 genes de *T. brucei* (2.63×10^6 codones) y 1721 genes de *T. cruzi* (9.86×10^5 codones). Los genes de *T. brucei* y *T. cruzi* fueron seleccionados al azar. Los datos obtenidos en cuanto al porcentaje de uso de codones en los Tritryps se presentan en la Tabla 6 (pag. 44). Al correlacionar estos valores con el número de isoceptores, se encontró una correlación positiva en *L. major* ($r_s = 0.80$) y *T. brucei* ($r_s = 0.64$), lo que significa que, al igual que en otros organismos, en tripanosomátidos el patrón de uso de codones está adaptado a la abundancia de los RNAt (Fig. 26A y 26B). En el caso de *T. cruzi*, el coeficiente de Spearman obtenido fue de $r_s = 0.35$ (Fig. 26C), lo que indica un bajo grado de correlación entre las dos variables, lo cual difiere de los otros dos tripanosomátidos. Este resultado puede ser consecuencia de que la cepa de *T. cruzi* CL Brenner que se utilizó para la secuenciación del genoma es un híbrido de dos cepas y algunos de sus genes están duplicados, mientras otros no.

Por lo anterior, se decidió repetir la prueba de Spearman graficando esta vez el uso de codones *versus* el número de genes de RNAt por aminoácido. Por ejemplo, en las tres especies, para alanina se graficó el uso de codones vs 5 genes totales de RNAt que pueden acarrear este aminoácido (Tabla 6). Esta vez, el coeficiente de Spearman fue más alto en *T. cruzi* ($r_s = 0.78$), lo que indica una correlación fuerte positiva entre ambas variables (Fig. 26F). Asimismo, se encontraron correlaciones positivas para *L. major* ($r_s = 0.84$) y *T. brucei* ($r_s = 0.85$) muy similares a las reportadas en el primer análisis (Fig. 26D y 26E).

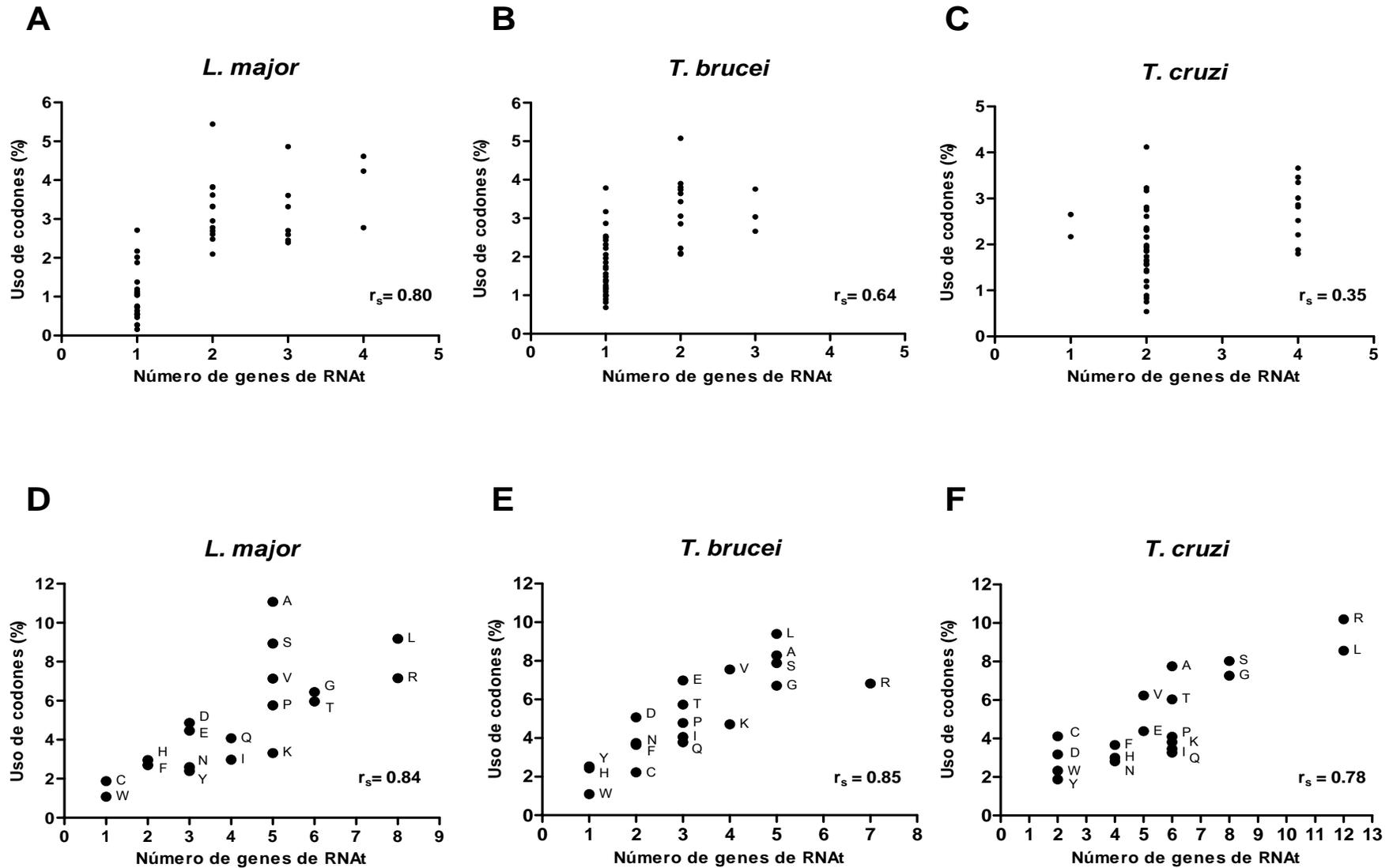


Figura 26. Correlación entre el número de genes de RNAt y el uso de codones en Trityps. Los paneles A-C muestran la correlación entre el uso de codones y el número de genes de RNAt por isoaceptor (45 tipos, no se incluye el RNAt-Sec) para las tres especies. Los paneles D-F muestran la correlación entre el uso de codones y el número de RNAt por aminoácido. Los coeficientes de correlación se evaluaron por la prueba de Spearman y sus valores se indican en cada panel (r_s), $p < 0.0001$ para los paneles A, B, D y E, $p < 0.02$ para el panel C, $p < 0.0001$ para el panel F.

5.12. Número de genes y estructura del RNAr 5S en tripanosomátidos

El análisis de las bases de datos de GeneDB de *L. major*, *T. brucei* y *T. cruzi* indica la presencia de 11, 9 y 135 genes de RNAr 5S, respectivamente. Es importante mencionar que en el caso de *L. major*, estos genes se encuentran localizados en los *clusters* que contienen RNAt y otros genes transcritos por Pol III, tal y como se mencionó en la sección 5.4 sobre la organización de los genes de RNAt (ver también Fig. 13). En el caso *T. brucei*, todos los genes anotados están localizados en tándem en el cromosoma 8. En *T. cruzi* ya ha sido reportado previamente (Hernández-Rivas *et al*, 1992) que éstos genes se encuentran organizados en tándem (repetidos de aproximadamente 500 pb); sin embargo, esto aún no se puede observar en las bases de datos porque los cromosomas todavía no se encuentran ensamblados en GeneDB.

Utilizando el programa Mfold plegamos las secuencias encontradas, encontrando que los RNAr 5S de los tripanosomátidos presentan la estructura típica de esa molécula, pues se aprecian los tallos y asas encontrados en esta molécula en otros eucariontes, es decir, las hélices I-V a forma de tallos, y las asas A-E. Dos de las asas, C y D, son estructuras *hairpin*, mientras que la B y E son asas que se encuentran internas en la estructura, mientras que la A funciona a manera de bisagra (Fig. 27).

El promotor de los genes de RNAt consiste en las cajas A y C y el elemento intermedio (EI), que se localizan dentro de la secuencia del gen (Fig. 2). Se identificaron los tres probables elementos de control interno en las secuencias de todos los genes anotados en GeneDB de RNAr 5S de *L. major* y *T. brucei* y de 10 genes tomados al azar de *T. cruzi*. Se obtuvieron secuencias consenso para las cajas A, C y el EI, tomando como referencia las secuencias *X. laevis* (Fig. 28).

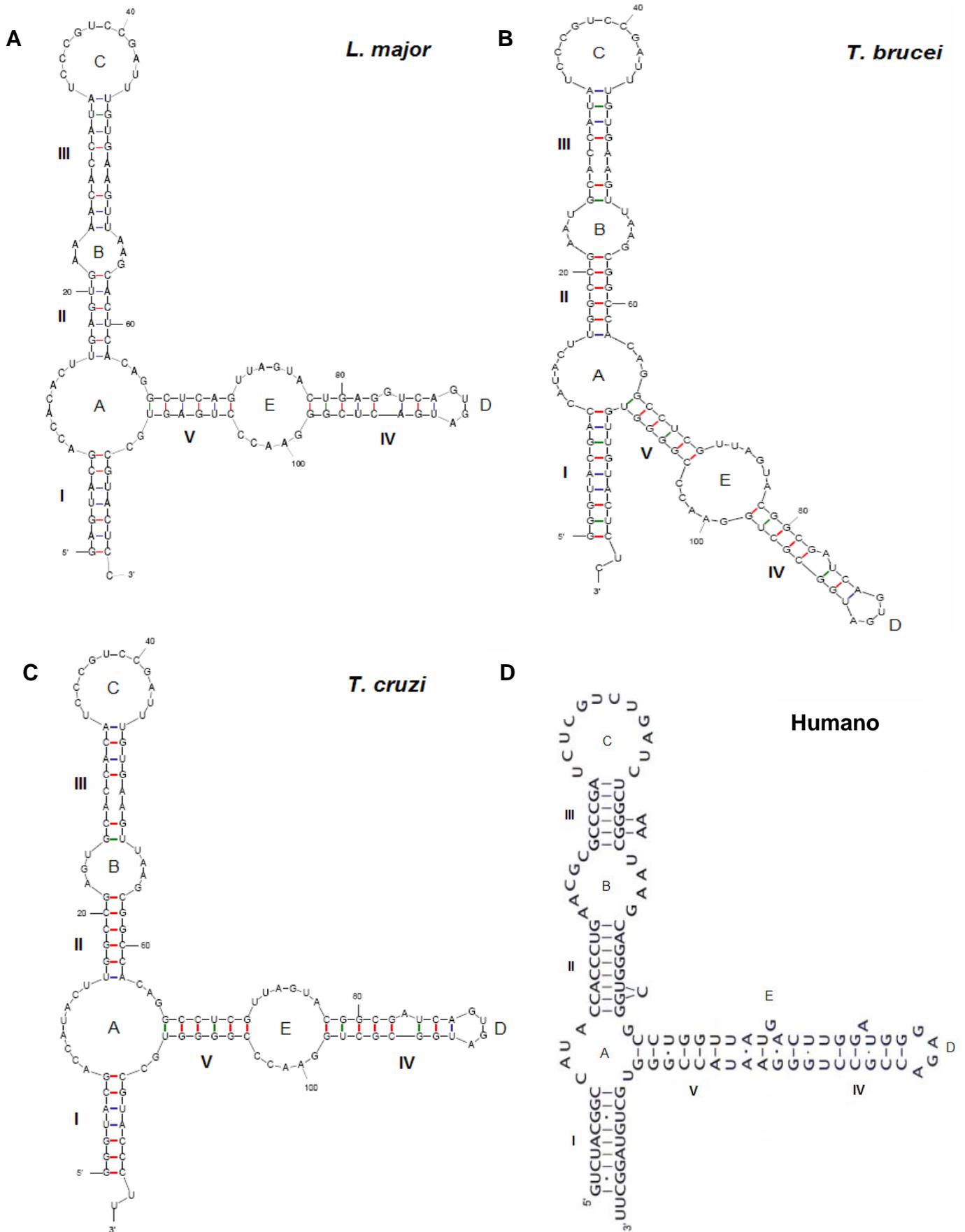


Figura 27. Estructuras secundarias de las secuencias de RNAr 5S en tripanosomátidos. A) RNAr 5S de *L. major*, B) RNAr 5S de *T. brucei*, C) RNAr 5S de *T. cruzi*, D) RNAr 5S de humano. Con números romanos se señalan las hélices I-V y con letras las asas A-E.

		CAJA A											
		1	2	3	4	5	6	7	8	9	10	11	12
<i>L. major</i>		A	A	G	T	T	A	A	G	C	A	C	T
<i>T. brucei</i>		A	A	G	T	T	A	A	G	C	G	G	C
<i>T. cruzi</i>		A	A	G	T	T	A	A	G	C	G	G	C
Tritryps		A	A	G	T	T	A	A	G	C	R	S	Y
<i>X. laevis</i>		A	A	G	C	C	A	A	G	C	A	G	G

		ELEMENTO INTERMEDIO						
		1	2	3	4	5	6	7
<i>L. major</i>		G	C	T	C	G/A	G	T
<i>T. brucei</i>		G	C	C	T	C	G	T
<i>T. cruzi</i>		G	C	C	T	C	G	T
Tritryps		G	C	Y	Y	N	G	T
<i>X. laevis</i>		G	C	C	T	G	G	T

		CAJA C											
		1	2	3	4	5	6	7	8	9	10	11	12
<i>L. major</i>		T	G	A	G	G	T	C	A	G	T	G	A
<i>T. brucei</i>		G	G	C	G	A	T	C	A	G	T	G	A
<i>T. cruzi</i>		G	G	C	G	A	T	C	A	G	T	G	A
Tritryps	G/T	G	G	A/C	G	R	T	C	A	G	T	G	A
<i>X. laevis</i>	T	G	G	G	G/A	T	G	G	G	A	G	A	C

Figura 28. Secuencias consenso de las cajas A, C y el Elemento Intermedio de los genes de RNAr 5S en Tritryps. Las secuencias consenso de cada elemento promotor se comparan contra la reportada para *Xenopus laevis*. Las posiciones conservadas entre tripanosomátidos y *X. laevis* se muestran en negritas. En caso de que dos nucleótidos sean posibles, se escribe primero el que ocurre con mayor frecuencia. R= purina G/A, Y= pirimidina T/C, S= G/C

6. DISCUSIÓN

Número de genes de RNA transferencia en tripanosomátidos

Nuestro análisis permitió la identificación de 82, 65 y 115 genes de RNA_t en los genomas de *L. major*, *T. brucei* (Fig. 13 y 14) y *T. cruzi*, respectivamente. Estos números son bajos cuando se comparan con otros organismos eucariontes, que generalmente contienen cientos de estos genes. Por ejemplo, *C. elegans* posee 568 genes, *Homo sapiens* 497 y *S. cerevisiae* 271, habiendo organismos como el pez cebra con ~6 000 genes (Marck y Grosjean, 2002 y Goodenbour *et al*, 2006). Una excepción es el microsporidio patógeno *Encephalitozoon cuniculi* que posee sólo 44 genes de RNA_t. Así, el bajo número de copias de RNA_t puede ser una característica distintiva de tripanosomátidos o incluso de algunos protozoarios, pues *Plasmodium falciparum* y *Giardia lamblia* poseen 44 y 63 genes de RNA_t, respectivamente (Morrison *et al*, 2007).

En *T. cruzi*, se identificaron tres genes de RNA_t que no correspondían al anticodón anotado en las bases de datos (Val-CAC) (Fig. 15). Estos genes se parecen sólo 61% al gen “real” de valina (el cual es 100% idéntico a su homólogo en *T. brucei* y 98% al de *L. major*), pero de manera interesante, se parece 75% al gen sRNA76 (anotado erróneamente como RNA_t-Sec en GeneDB) de *T. brucei*, lo que sugiere que los genes de RNA_t mal anotados podrían representar el homólogo de dicho transcrito en *T. cruzi*. El sRNA76 es una RNA no codificante que coprecipita con el RNA 7SL en *T. brucei* (Béjà, 1993), y que se ha encontrado también en otros cinetoplástidos como *Leptomonas collosoma* (Liu *et al*, 2003); en ambos casos, se trata de un RNA_t-like que posee cajas promotoras similares a las de los RNA_t funcionales, secuencias que se encuentran presentes en los RNA_t mal anotados de *T. cruzi*. Otra alternativa es que estas secuencias correspondan a RNA_t con isotipo (anticodón) indeterminado, los cuales están presentes en organismos como *C. elegans*, levaduras y humano (<http://lowelab.ucsc.edu/tRNAscan-SE>). A la fecha se desconoce la función de los RNA_t con isotipo indeterminado.

Además de los tres RNAt-Val mal anotados en las bases de datos, también se identificó el gen Tc00.1047053507579.16, anotado como un gen de RNAt-Ile (Fig. 16). Sin embargo, tRNAscan no lo reconoce como un RNAt (el anticodón no se encuentra bien definido en la estructura, posee inserciones en la secuencia y arroja un *score* < 30) y parece tratarse más bien de un pseudogen, pues se parece sólo 22% al resto de los genes de RNAt-Ile anotados y no posee en el extremo 3' la señal de término de la transcripción. Los pseudogenes de RNAt se han identificado tanto en bacterias como en eucariontes y se trata de genes con mutaciones que impiden el proceso de maduración. En humano se ha identificado actividad transcripcional de algunos pseudogenes *in vivo*, aunque su función aún no se ha dilucidado (Mans, *et al*, 1991).

Otra posibilidad es que estos genes de RNAt en *T. cruzi* correspondan a SINEs (*short interspersed elements*); pues algunos de éstos son retrotransposones derivados de RNAt, encontrados en los genomas eucariontes, que no codifican una transcriptasa reversa, por lo que dependen de otros retrotransposones para moverse. Algunos SINES aún poseen las cajas A y B, (Wickstead *et al*, 2003). Recientemente, se ha reportado una clase de SINES en *Chlamydomonas reinhardtii* (Cognat, *et al*, 2008) que guardan una estructura altamente similar a la de los RNAt de ácido aspártico y arginina, con presencia de cajas A y B, tal y como sucede con los RNAt de isotipo indeterminado que encontramos en *T. cruzi*. Esta probabilidad está apoyada por el hecho de que los retroelementos son muy comunes en *T. cruzi*, pues representan ~5% del genoma (El-Sayed *et al*, 2005). En *L. major* no se encontraron RNAt de isotipo indeterminado y, curiosamente, no se han encontrado retroelementos en este organismo. La función específica de los SINES aún no se ha entendido completamente, pero se sabe tienen un impacto importante en la estructura, evolución y expresión del genoma y en mamíferos se sabe que son capaces de regular la transcripción y traducción (Cognat *et al*, 2008). Su función en tripanosomátidos es aún menos entendida.

RNAt isoceptores

El análisis de los anticodones de los RNAt en tripanosomátidos reveló la presencia de 46 especies isoceptoras en cada una de las tres especies, 45 isotipos que corresponden a los aminoácidos clásicos (reportados previamente por Horn, 2008) y uno que corresponde al RNAt selenocisteína (Tabla 6) (cabe recordar que el anticodón de dicho aminoácido es TCA, una señal utilizada como codón de paro). Este número de isoceptores en los *Trityps* es similar a otros organismos eucariontes, que poseen de 41 a 55 isoceptores, según la especie (Goodenbourg y Pan, 2006, Marck y Grosjean, 2002). Dieciséis anticodones no fueron encontrados en el genoma de los tripanosomátidos (Tabla 6, marcados en rojo) y, sin embargo, sus codones respectivos sí se encuentran presentes en las proteínas de estos organismos (Tabla 6, datos sobre uso de codones) (Horn, 2008). Por ejemplo, el anticodón ATG de los RNAt-His no se encuentra presente en los tripanosomátidos, pero el codón complementario CAT sí se encuentra en las proteínas. Es posible leer estos codones sin los dieciséis anticodones faltantes gracias a los apareamientos flexibles entre la tercera posición del codón y la primera del anticodón (Fig. 17), lo que permite que algunos isoceptores reconozcan dos o más codones que difieran sólo en la tercera base.

Al analizar los anticodones reportados en la Tabla 6, se observó que para un cierto aminoácido sólo existe el anticodón ya sea con G1 ó con A1 (A es convertida a I en el RNAt maduro), el cual es utilizado para leer los codones con C3 y U3; esto significa que los tripanosomátidos utilizan una estrategia de codificación de recambio $A1 \leftrightarrow G1$. Este modo de codificación es el que se encuentra en el genoma de otros eucariontes como *C. elegans*, *Homo sapiens* y *Arabidopsis thaliana*. Al igual que en estos eucariontes, entre los tripanosomátidos los RNAt con A1 en el anticodón sólo se encuentran en las familias de cuatro codones: leucina, valina, serina, prolina, treonina, alanina y arginina (excepto en el caso de glicina), así como en isoleucina, familia de tres codones. En estos casos, A1 lee C3 y U3 en los codones complementarios. Las familias de codones que presentan sólo el anticodón con G1 son aquellas de dos codones:

asparagina, ácido aspártico, cisteína, histidina, fenilalanina, serina (la familia de dos codones), tirosina y glicina, única familia de cuatro codones que presenta sólo el anticodón con G1. En otros organismos, 2 estrategias más de decodificación de codones son posibles: la estrategia de reemplazo de C1 y la de recambio A1, G1 y C1. En la primera, reemplazo de C1, además de los anticodones que comienzan con A1 ó G1, no se encuentran en el genoma algunos anticodones con C en la primera posición del anticodón, por lo que los anticodones con U1 leen los codones con A3 y G3 correspondientes. Eucariontes como levaduras, *Drosophila* y *E. cuniculi* utilizan esta estrategia de decodificación, en la que el número de isoaceptores en el genoma puede ser reducido hasta 33. La última estrategia, el recambio o reemplazo de A1, G1 y C1 es usada sólo por bacterias, en la que los anticodones con U1 leen todos los codones sinónimos para un cierto aminoácido; es decir, U1 lee los codones con U3, G3, A3 y C3, por lo que el número de isoaceptores por genoma puede ser reducido hasta 30 (Marck y Grosjean, 2002).

RNA_t de Metionina elongador e iniciador

Los RNA_t-Met de elongación e iniciación se identificaron en los genomas de *L. major*, *T. brucei* y *T. cruzi*. Cabe mencionar que un RNA_t de cada isotipo ha sido descrito previamente en *T. brucei* por Tan *et al* (2002). Estos dos tipos de RNA_t-Met poseen características específicas que los distinguen entre sí, la mayoría de las cuales se encontraron en las estructuras de los tripanosomátidos. Uno de los principales determinantes para distinguir entre el elongador y el iniciador es el par A1:U72 que está presente en todos los iniciadores (A1:U71 en tripanosomátidos, Fig. 18A) (Farrugio *et al*, 1996), mientras que los elongadores presentan G:C en esa posición (Fig. 18B). En humano y levadura, esta posición es determinante para evitar que el RNA_t iniciador participe durante el proceso de elongación y al parecer es necesaria para que el RNA_t se una al factor de iniciación 2 de eucariontes (eIF2); mutaciones en esta posición permiten a Met-i actuar como Met-e, debido a su unión con el factor de elongación eEF2 (Drabkin *et al*, 1998). En levadura, se determinó que el par A1:U72 es importante para la unión del iniciador a la subunidad 40S

ribosomal (Kapp *et al*, 2006), primer paso durante la iniciación de la síntesis de proteínas. Es probable que este par A1:U72 tenga funciones muy similares en tripanosomátidos.

Otra característica distintiva entre los dos tipos de RNAt-Met es la presencia de tres pares G:C consecutivos al final del tallo del anticodón en los RNAt iniciadores conservados tanto en eucariontes como en bacterias. En *E. coli*, estas bases son indispensables para iniciar de manera eficiente la traducción al favorecer la interacción del RNAt con el sitio P del ribosoma (Mandal *et al*, 1996). En humanos, mutaciones en estas posiciones reducen también la eficiencia de la traducción (Drabkin *et al*, 1993); sin embargo, en levadura las mismas mutaciones apenas afectan el crecimiento celular (Drabkin *et al*, 1998). En levadura, las G:C consecutivas en el RNAt-i parecen requerirse para una interacción adecuada con factores de iniciación (eIF2) y con la subunidad 40S (Kapp *et al*, 2006). Sorpresivamente, en tripanosomátidos se encontraron los tres pares consecutivos G:C tanto en el iniciador como en el elongador (Fig. 18), por lo que estas bases no son un determinante para distinguir entre ambos RNAt. Así, los tripanosomátidos podrían utilizar otras características del RNAt-i para dirigirlo al sitio P del ribosoma y para interactuar con los factores de iniciación y con sus subunidades ribosomales. Este trabajo puede servir como punto de partida para el diseño de estrategias experimentales que permitieran comprender mejor cuáles son los efectos reales de estos pares de bases G:C y de otras características en el reconocimiento del RNAt-i en tripanosomátidos.

Otros determinantes para la función de iniciador que se encontraron conservadas entre tripanosomátidos son A53, A58 y A59 en el asa T ψ C, mientras que en el elongador se encontró U54, U59 y U60 en las posiciones equivalentes, tal y como se ha reportado para otros eucariontes. Estas posiciones del asa T ψ C son también primordiales para evitar que Met-i participe durante la elongación. En *S. cerevisiae*, las mutaciones en las posiciones A54 y A60 en el iniciador no afectan su función pero sí el crecimiento celular (von Pawel-Rammingen *et al*, 1992); y por experimentos con alelos nulos del Met-e se demostró que U54 es crucial para la actividad de elongador (Aström *et*

al, 1993). Se ha sugerido que el factor eEF1 contacta estos puntos en la región T ψ C (Drabkin *et al*, 1998) y que estas posiciones son importantes para la unión de las subunidades ribosomales y la estabilidad de ese complejo (Kapp *et al*, 2006). En *Trityps*, las posiciones A54 y A60 en el iniciador y sus correspondientes en el elongador podrían tener papeles similares (Fig. 18).

Otros determinantes para distinguir entre Met-i y Met-e son el par C3:G69 (G3:C70 en Met-e), las bases A19 y A20 (G19 y C20 en el elongador) y las asas D y la del anticodón formadas por 7 nt, mientras que en el elongador están formadas por 8 y 9 nt, respectivamente), características encontradas en los RNAt-Met de los *Trityps* (Fig. 18).

Organización de los genes de RNAt

Los genes de RNAt en *L. major* y *T. brucei* están organizados en *clusters* de 2 a 10 genes repartidos en varios cromosomas (Fig. 13 y 14, respectivamente, Tabla 8), y en *L. major* se encuentran también colindando con otros genes transcritos por Pol III, como genes RNAr 5S y RNA pequeños nucleares. Dentro de los *clusters*, los genes pueden estar orientados tanto en la cadena codificante como en la hebra no codificante del DNA. Entre eucariontes, los genes de RNAt se encuentran generalmente distribuidos al azar en el genoma, aunque esta distribución depende del organismo. En humano, la distribución parece ser no azarosa, ya que más del 25% de los genes de RNAt se localizan en una región de ~4 Mb en el cromosoma 6, una región que representa sólo 0.1% del genoma humano, pero contiene un conjunto casi completo de genes de RNAt. En *L. major*, la distribución de los genes de RNAt no parece ser totalmente al azar, pues casi la mitad de los cromosomas no contienen RNAs de transferencia. Además, 60 genes de RNAt (72%) se localizan sólo en 7 cromosomas (9, 11, 23, 24, 31, 34 y 36) lo que representa ~26% del genoma. En *T. brucei*, 40 (61.5%) de los genes de RNAt se encuentran en sólo 3 cromosomas (4, 7, 8), que representa ~24% del genoma.

En *S. cerevisiae*, aunque los genes de RNAt se encuentren dispersos en todo el genoma, éstos se agrupan en el nucleolo junto con genes de RNAr 5S, lo que podría facilitar el acceso a factores de transcripción comunes, así como a enzimas de maduración del pre-RNAt que también se localizan en nucleolo (Thompson *et al*, 2003). Aún no se sabe si los genes de RNAt en tripanosomátidos tienen un punto específico de localización celular como en levadura.

En *S. pombe* y *C. elegans*, algunos genes de RNAt se encuentran agrupados en centrómeros y actúan como barreras de dispersión de la cromatina (Scott *et al* 2007). Probables regiones centrómicas ha sido reportadas en algunos cromosomas de *T. brucei* y *T. cruzi* (Obado *et al*, 2007), que colocan con regiones de cambio de hebra que separan DGCs divergentes. Estas regiones no parecen tener genes de RNAt, sin embargo, dos *clusters* de genes de RNAt en *L. major* (Chr09-II y Chr10) (ver Fig 13, tabla 8), y un *cluster* en *T. brucei* (Chr04-I) (Fig. 14, Tabla 8) se localizan en áreas del genoma que probablemente correspondan a regiones centroméricas (áreas de cambio de hebra que separan genes divergentes de Pol II). Por lo tanto, es posible que en tripanosomátidos, al igual que en *S. pombe* y *C. elegans*, algunos genes de RNAt puedan contribuir a la actividad centromérica. Los genes de RNAt se encontraron separados por distancias intergénicas cortas, de 35 a 5406 bases y de 43 a 3172 bases en *L. major* y *T. brucei*, respectivamente (Fig. 13 y 14), arreglos que se encuentran en otros eucariontes. Por ejemplo, en *S. pombe*, aunque la mayoría de sus RNAt están dispersos en el genoma, posee 4 copias de RNAt-Met separados por 5 o 7 bases de otros RNAt de Asp o Ser (Marck y Grosjean, 2002) pero sus distancias intergénicas pueden ser también mayores, como en otro *locus*, hay un RNAt-Ile separado de su gen de RNAt-Glu adyacente por 2786 pb.

Relación espacial entre los genes de Pol II y Pol III en tripanosomátidos.

Al analizar los mapas cromosomales reportados para *L. major* se identificaron 39 regiones de cambio de hebra convergentes en el genoma. Catorce de ellas (37%) contienen al menos 1 gen de RNAt. Entre estas regiones se encuentra la localizada en el cromosoma 3 (ver fig. 13), que contiene un gen de tRNA-Lys que participa en el término de la transcripción (de Pol II) de las dos unidades de transcripción convergentes presentes en el cromosoma (Martínez-Calvillo *et al*, 2004). Es probable que los genes de tRNA localizados en los otros 13 DGCs convergentes también participen en el término de la transcripción de Pol II. En *S. cerevisiae*, cuando la transcripción de dos genes convergentes no termina de manera apropiada (al final de cada gen) la expresión de ambos genes se ve drásticamente afectada por un mecanismo conocido como colisión transcripcional (Prescott y Proudfoot, 2002). Se presume que la presencia de señales de término de la transcripción entre DGCs convergentes en los tripanosomátidos evitaría un proceso similar a la colisión transcripcional. En el caso de los DGCs convergentes que no contienen tRNAs en la región de cambio de hebra, no se descarta la presencia de otros genes transcritos por Pol III aún no identificados.

En términos del número total de *clusters* de RNAt presentes en el genoma de *L. major*, 14 *clusters* de 31 (45%) se encuentran entre DGCs convergentes (ej. Chr03 en figura 13 y tablas 8 y 9). En *T. brucei* se encontró algo similar, pues 36% (9/25) de los *clusters* de RNAt se encuentran entre DGCs convergentes (ej. *cluster* Chr04-III en figura 14 y tabla 8). Aun en el caso de *clusters* de RNAt localizados dentro de unidades direccionales de genes (DGCs) (por ejemplo, *cluster* Chr11-I, figura 13 y tablas 8 y 9) es probable que los *clusters* de RNAt interrumpen la transcripción por Pol II. Es decir, que la transcripción de la parte “izquierda” del DGC termine antes del *cluster* de genes de RNAt y se reanude cerca del primer gen en la parte “derecha” del DGC, después de los RNAt. Las regiones intergénicas entre los genes codificadores de proteínas y los *clusters* de genes de RNAt son generalmente largas, lo que podrían proporcionar suficiente espacio para el inicio y término de la transcripción por Pol II y por Pol III.

Sintenia entre genes de Pol III

La mayoría de los *clusters* de RNAt no muestran sintenia entre las tres especies de tripanosomátidos, ya que sólo 4 *clusters* mostraron conservación en el orden de los genes (Fig. 19). Esto es diferente a lo observado con los genes transcritos por Pol II, ya que los genomas de los tripanosomátidos son altamente sinténicos entre sí para estos genes; por ejemplo, *T. brucei* y *L. major* presentan 110 bloques de sintenia.

Tal cual se encontró en *Trityps*, entre varias especies de oomicetos del género *Phytophthora* se ha observado poca sintenia en genes transcritos por Pol III, a pesar de que los genes codificadores de proteínas son altamente sinténicos (Tripathy y Tyler, 2006). Esto sugiere que la conservación en el orden de genes de RNAt es funcionalmente menos importante que la conservación en el orden de los genes codificadores de proteínas. Por otra parte, en *Plasmodium* spp. se ha observado que los genes de RNAr son no sinténicos aún cuando los genes codificadores de proteínas presentan alto grado de sintenia entre especies de ese género. Puesto que los genes de unidades ribosomales (18S-5.8S-28S) representan puntos de ruptura de sintenia entre los genomas de *P. falciparum* y *P. y. yoelii*, se ha sugerido que estos sitios son puntos de ruptura de cromosomas y recombinación que contribuyeron a la evolución del género *Plasmodium* (Carlton *et al*, 2005). Ya que en general no se encontró sintenia entre genes de RNAt, sería muy interesante investigar en un futuro si estos genes pudieran haber contribuido a la evolución del genoma de los tripanosomátidos de una manera similar a los RNAr de *Plasmodium*. Además, sería también interesante investigar si los genes de los cuatro *clusters* de RNAt sinténicos entre los *Trityps* son, por alguna razón evolutiva, transcripcionalmente más activos que los *clusters* no sinténicos.

Intrones en los genes de RNAt en tripanosomátidos.

En tripanosomátidos, al igual que en otros eucariontes, se encontró un intrón en el RNAt de tirosina (GTA); éste había sido previamente reportado en *T. brucei* (Schneider *et al*, 1993). Tal y como sucede generalmente en eucariontes, los intrones en tripanosomátidos se encontraron exactamente 1 nt después del anticodón, entre las bases 37 y 38 (Fig. 20). Aunque no hay una regla clara acerca del tamaño del intrón, entre otros eucariontes como levaduras, *C. elegans*, humano, *A. thaliana* y *P. falciparum* el intrón del RNAt-Tyr (GTA) va de 10 a 20 pb (Marck y Grosjean, 2002; <http://lowelab.ucsc.edu/tRNAscan-SE/>). Esto es similar a lo encontrado en Trityps, pues el intrón en *L. major* y *T. brucei* es de 11 pb, mientras que en *T. cruzi* es de 13 nt (Fig. 20). A diferencia de los tripanosomátidos, la presencia de intrones en los RNAt de diversos isotipos es común entre otros eucariontes. En general, no hay una regla que dicte en que isoaceptores se encuentran los intrones, sin embargo se les observa con frecuencia en los RNAt de Ile y Leu. Entre eucariontes, los intrones van de 7 a 60 pb de longitud. Asimismo, *S. cerevisiae* y *S. pombe* son los eucariontes que presentan un mayor número de isoaceptores con intrones: 10 y 14, respectivamente, muchos más que en tripanosomátidos (1 isoceptor con intrón) (Marck y Grosjean, 2002)

Se ha sugerido que los intrones de los RNAt en eucariontes tienen un papel importante en el procesamiento del pre-RNAt: en la conversión de uridina a pseudouridinas o bien, que el intrón sea importante para el reconocimiento por las endonucleasas que participan en el *splicing* del mismo intrón (Schneider *et al*, 1993). Así, en Trityps los intrones son raros no sólo en genes codificadores de proteínas, sino también en RNAt.

Secuencias consenso de los elementos promotores

Se identificaron los probables elementos de control interno, las cajas A y B, en todos los genes de RNAt en los tres tripanosomátidos. La secuencia de estos elementos es similar a la de otros eucariontes como *S. cerevisiae* (Fig. 21). Para obtener la secuencia consenso tuvimos que dividir los genes de RNAt en dos clases: clase I, aquellos RNAt con un asa variable corta y clase II, aquellos con un brazo variable largo (grupo que comprende únicamente los RNAt de leucina y serina). Esta clasificación de los genes nos permitió obtener una secuencia consenso más conservada para cada uno de los grupos, lo cual no sucedía si se consideraban todos los genes como un solo grupo. No tenemos conocimiento de que en algún otro organismo se hayan analizado por separado las secuencias de los elementos de control interno de los tRNA de las clases I y II.

La mitad de las bases de la secuencia consenso de la caja A (posiciones 1, 2, 4, 7, 10 y 11) son idénticas entre los genes de clase I y clase II (e idénticas a la secuencia consenso de levadura) (Fig. 21). Sin embargo, la posición 5 es diferente entre ambas clases ya que los genes de clase I pueden tener cualquier nucleótido, mientras que los genes clase II tienen siempre C. Asimismo, los genes de clase I presentan C o T en la posición 6, mientras que los de clase II tienen A o G. Con respecto a la caja B, la posición 8 es diferente en ambas clases: una A está siempre presente en los genes de clase II, mientras que los de clase I pueden tener cualquier nucleótido. Alrededor del 20% de los genes de RNAt clase I en *Trityps* tienen un nucleótido adicional en la caja A, entre las posiciones 9 y 10; en lo que respecta a los genes de clase II, los RNAt de leucina son los que poseen una T entre esas posiciones de la caja A.

En ningún otro organismo se ha reportado que existan diferencias en las secuencias de las cajas A y B entre genes de RNAt clase I y clase II. De hecho, todas las secuencias consenso reportadas hasta ahora incluyen ambas clases de genes, por lo que las secuencias parecen ser menos conservadas de lo que realmente son. Haciendo un análisis de las secuencias de las cajas

promotoras de algunos genes de RNAt en *S. cerevisiae* y separando los genes en clase I y II, observamos que las secuencias consenso son más conservadas para genes de la misma clase y que se aprecian diferencias con respecto a las secuencias consenso reportadas (que incluyen todos los genes de RNAt por igual) (datos no mostrados). Esto sugiere que las diferencias en las secuencias de las cajas A y B entre genes de RNAt clase I y II no son exclusivas de los Trityps, lo cual podría tener un significado funcional para los eucariontes en general.

También obtuvimos la secuencia consenso de los elementos promotores de los genes de RNAt-Sec en Trityps. Al igual que en otros eucariontes, la caja A posee 1 nt extra dentro de la secuencia (Fig. 22). Con respecto a la caja B, encontramos dos cambios en los Trityps (en la primera y última bases) que no están presentes en otros organismos. En *Xenopus* se observó que la caja A no es esencial para la transcripción del gen, pues al eliminarla no se disminuye la transcripción del mismo, aunque el transcrito producido es de menor tamaño. Por su parte, la caja B tampoco parece ser esencial, pero sí incrementa la eficiencia de la transcripción del gen, pues mutaciones en la caja B causan una reducción en el nivel de transcripción; y similar a la caja A, se observan transcritos mayores y menores al tamaño esperado. Por esto, las cajas A y B podrían ser necesarias para una correcta iniciación de la transcripción y/o para un procesamiento adecuado de los pre-RNAt-Sec (Carbon y Krol, 1991; Park *et al*, 1995). En Trityps, las cajas A y B del RNAt-Sec podrían tener un efecto similar en la transcripción del gen. Otra característica distintiva de los genes de RNAt-Sec son sus elementos promotores extragénicos, río arriba del SIT y esenciales para su transcripción. Éstos están conformados por una caja TATA, un elemento proximal y un elemento activador, elementos que se encuentran también en los promotores de los genes de RNAsn U6 (Hubert *et al*, 1996; Myslinski *et al*, 1993).

Al igual que en *Xenopus*, es muy probable que en tripanosomátidos existan elementos río arriba del SIT que tengan una influencia determinante en la transcripción del RNAt-Sec. Se ha hecho referencia a los genes de RNAt-Sec de otros eucariontes como un híbrido, en cuando a sus

regiones promotoras, por poseer tanto elementos similares a los RNAt clásicos (cajas A y B) como elementos extragénicos de un promotor tipo III (caja TATA y PSE). Sería interesante probar experimentalmente cuáles serían, en su caso, los elementos promotores extragénicos del RNAt-Sec en tripanosomátidos. En tripanosomátidos la expresión del RNApn U6 es atípica en que requiere de las cajas A y B de un RNAt situado río arriba.

Señales de término de la transcripción en genes de RNAt

En *L. major*, se ha demostrado que la transcripción del gen de RNAt localizado en el cromosoma 3 termina dentro de un tracto de cuatro Ts consecutivas (Martínez-Calvillo *et al*, 2004). Se observó que *L. major* y *T. brucei* los genes de RNAt poseen en promedio 4 Ts (en un rango de 4-10) para terminar la transcripción, mientras que *T. cruzi* tienen en promedio 6-7 bases (con un rango de 4-16), diferencia que es estadísticamente significativa (Fig. 23A). En *S. cerevisiae* y *S. pombe* se requieren primordialmente 6 y 7 nt, respectivamente, para el término de la transcripción, mientras que en humano y ratón son más frecuentes las secuencias de 4 nt (Braglia *et al*, 2005). Así, la longitud de la secuencia de Ts parece ser específica de especie aún entre especies filogenéticamente relacionadas.

La presencia de una segunda secuencia, o *back up*, alrededor de 30 pb río abajo de la primera es variable entre tripanosomátidos. El porcentaje de *T. cruzi* (48%) es muy similar al de otros eucariontes como *S. cerevisiae* (44%), *S. pombe* (53%). Por otra parte, los porcentajes de *L. major* y *T. brucei* (16% y 27%, respectivamente) son menores que estos organismos, incluyendo mamíferos (ratón y humano poseen 33% y 31% de genes con *back up*, respectivamente) (Fig. 24).

Así, *L. major* y *T. brucei* poseen una secuencia corta de Ts que parece ser suficiente para lograr una terminación correcta de la transcripción en la mayoría de los genes de RNAt, mientras que *T. cruzi* parece requerir secuencias más largas de T, además de una segunda secuencia (*back-*

up) para terminar la transcripción de los RNAt. Esto indica que el mecanismo de la terminación de la transcripción por Pol III en *L. major* y *T. brucei* podría ser diferente a *T. cruzi* y a otros eucariontes. La poca frecuencia de señales *back up* en *L. major* y *T. brucei* podría sugerir que sus señales de término son fuertes, pues las secuencias río arriba y río abajo de la secuencia de Ts pueden contribuir a incrementar o disminuir la fuerza de la señal de término, tal y como se ha observado en *S. cerevisiae* (Braglia *et al*, 2005). Sería interesante realizar experimentos para caracterizar mejor estas regiones de término y saber qué bases o secuencias (además de las regiones de Ts) debilitan o potencializan la terminación de la transcripción en Trityps. Aunque en tripanosomátidos los genes de RNAt se encuentran agrupados en *clusters*, la presencia de la secuencia de Ts localizada río abajo de todos los DNAt sugiere que éstos son transcritos como monocistrones, lo que es común entre eucariontes.

RNAt isodecodificadores.

En *L. major* y *T. brucei* se localizó sólo un isodecodificador (Ser-GCT y Ala-CGC, respectivamente) que se parece entre 97-98% a al isoceptor principal (Fig. 25, tabla 11). En el caso de *L. major*, la diferencia entre las secuencias (sólo 1 nt) se encuentra muy cerca de la caja A. De igual manera, en *T. brucei*, una de las diferencias se encuentra cerca de la caja A, mientras que la otra reside dentro de la caja B, en una base altamente conservada (la posición 5 de la secuencia consenso). En *T. cruzi*, se encontraron 6 isodecodificadores, 3 de los cuales (Glu, Arg, His, Ala-TGC) difieren del isoceptor mayoritario por bases que se encuentran dentro de la caja B, en nucleótidos variables (Fig. 25, tabla 11). Como en Trityps, en humano la mayoría de las variaciones en la secuencia entre los isodecodificadores se localizan dentro de las regiones promotoras internas, con más frecuencia en la caja B (Goodenbour y Pan, 2006). Así, las variaciones en las secuencias dentro de los elementos de control interno en los genes de RNAt isodecodificadores sugieren que en tripanosomátidos, como en otras especies, podría existir una regulación diferencial de la transcripción por Pol III. El hecho de que una C altamente conservada de la posición 10 de la

caja B en el gen TbtRNA-Ala.02 se haya cambiado a una T en el isodecodificador correspondiente (TbtRNA-Ala.01) apoya esta posibilidad. En humano se ha observado que la expresión de RNAt varía entre tejidos (Dittmar *et al*, 2006; Goodenbour y Pan, 2006), por lo que los RNAt podrían tener alguna función específica de tejido o en el desarrollo y/o diferenciación celular. Resultaría muy interesante determinar si los RNAt isodecodificadores de tripanosomátidos se expresan diferencialmente.

Uso de codones en tripanosomátidos

En tripanosomátidos se ha reportado que la tendencia a utilizar ciertos codones correlaciona con el número de copias génicas de RNAt y con el nivel de expresión proteica (Horn, 2008). Este análisis se realizó haciendo un análisis de 60 000 codones de genes codificadores de proteínas altamente expresadas en los tres parásitos. Nosotros quisimos realizar un análisis similar, pero incluyendo todos los genes de *L. major* y un número muy elevado de genes de *T. brucei* y *T. cruzi* seleccionados al azar. El coeficiente de correlación obtenido para *L. major* y *T. brucei* sugiere que existe una correlación positiva entre el número de genes de RNAt y el uso de codones (Fig. 26 A-B). Esto es similar a lo encontrado en otros organismos, tanto procariontes como eucariontes (Duret, 2000; Kanaya *et al*, 1999; Percudani *et al*, 1997; Rocha, 2004). Esto significa que los codones más favorables son los correspondientes al RNAt más abundante y eficiente. Esto se debe a que la traducción es un proceso altamente demandante de energía, por lo que los genes relacionados con el proceso se encuentran bajo presión selectiva. En *Escherichia coli*, se ha encontrado que la correlación es más fuerte a mayores velocidades de crecimiento, lo que indica la importancia de estas variables en la eficiencia de traducción (Rocha, 2004). Todavía no se ha entendido completamente cómo un anticodón reconoce de manera preferencial cierto codón en particular, aun cuando sea capaz de reconocer más, y cómo el contenido de genes de RNAt co-evolucionan con el uso de codones para lograr una optimización de la maquinaria de traducción (Kanaya *et al*, 2001; Rocha, 2004).

En *T. cruzi* no se obtuvo una correlación positiva entre las variables mencionadas (Fig. 26C), probablemente porque la cepa de este organismo utilizada para el proyecto de secuenciación es un híbrido de dos cepas y algunos de sus genes podrían estar duplicados mientras que otros no. La correlación positiva para este organismo se observó sólo cuando se graficó el número total de RNAt por aminoácido *versus* el uso de codones ($r_s = 0.74$, Fig. 26F), mientras que el resultado se mantuvo para los otros dos organismos, pues se encontró $r_s = 0.84$ y $r_s = 0.85$ para *L. major* y *T. brucei*, respectivamente (Fig. 26D-E); estos coeficientes nos llevan a la misma conclusión del primer análisis, que existe una correlación entre el número de genes de RNAt y el uso de codones en estos parásitos.

RNA ribosomal 5S en tripanosomátidos

Con respecto al RNAr 5S, en las bases de datos de los genomas de *L. major*, *T. brucei* y *T. cruzi*, se encuentran anotados 11, 9 y 135 genes, respectivamente. Tal como ocurre con el número de genes de RNAt, el número de genes del RNAr 5S anotado en GeneDB en *L. major* y *T. brucei* es muy bajo, considerando que la mayoría de los organismos contienen de 100 a 2000 copias del gen. En *L. major*, los 11 genes del RNAr 5S no se encuentran organizados en *cluster*, sino que están dispersos en el genoma, asociados a otros genes transcritos por Pol III. En este organismo se han identificado 6 copias de la unidad de transcripción del RNAr en el cromosoma 27, en arreglo tipo tándem (Martínez-Calvillo, S. *et al*, 2001). Por lo tanto, el número de genes del RNAr 5S y de la unidad del RNAr es similar en *L. major*. En lo que respecta a *T. brucei*, hay también correspondencia entre el número de genes de RNAr 5S (9 copias en tándem) y el número de unidades del RNAr anotadas (8 copias completas, más varias fragmentadas). En lo que toca a *T. cruzi*, la falta de ensamblaje de los *contigs* secuenciados dificulta el análisis de las unidades del RNAr; aunque el número de genes individuales anotados indica la presencia de ~8 copias de dicha unidad. Por lo tanto, en *T. cruzi*, el número de genes del RNAr 5S sería ~16 veces mayor al número

de copias de la unidad del RNAr. A este respecto, se ha reportado en una cepa mexicana de *T. cruzi* (la cepa del proyecto del genoma de *T. cruzi* es brasileña) la presencia de ~1600 genes del RNAr 5S, número 16 veces mayor que el de las unidades del RNAr (Hernández-Rivas *et al*, 1992). De igual manera, en una cepa de *T. brucei* diferente a la que se usó en el proyecto del genoma de dicho parásito se reportó la presencia de ~1500 genes del RNAr 5S (Hasan *et al*, 1984). Esto sugeriría un elevado polimorfismo en cuanto al número de genes del RNAr 5S entre diferentes cepas de *T. cruzi* y *T. brucei*. Sin embargo, subestimaciones y/o sobreestimaciones del número de genes no pueden descartarse.

Se identificaron los probables elementos de control interno (la caja A, la caja C y el Elemento Intermedio) en todos los genes de RNAr 5S de *L. major* y *T. brucei* y 10 genes tomados al azar en *T. cruzi*. Estas secuencias se aproximan a las secuencias promotoras de *X. laevis* (Fig. 28). La mayoría de las bases de la secuencia consenso de la caja A (posiciones 1, 2, 3, 6, 7, 8 y 9 en la figura 28) son idénticas entre tripanosomátidos y la secuencia de *X. laevis*. Las posiciones 4 y 5 y 10 son altamente conservadas entre tripanosomátidos y aunque a simple vista parecería que varían de la secuencia de *Xenopus*, el hecho es que en las posiciones 4 y 5 se encuentra siempre una pirimidina (T o C), mientras que en la posición 10 se encuentra siempre una purina (G o A).

Acerca del EI, 4 bases son conservadas entre tripanosomátidos y entre tripanosomátidos y *Xenopus*: las posiciones 1, 2, 6 y 7. De manera similar a la caja A, en las posiciones 3 y 4 se encuentra una preferencia por pirimidinas C o T, tanto en *Trypanosoma* como en *Xenopus*. Finalmente, en la caja C las posiciones conservadas entre tripanosomátidos son la 2, 4 y 6-12, es decir, es altamente conservada entre tripanosomátidos. Sin embargo, al comparar con la secuencia de *Xenopus*, la similitud es muy baja, pues sólo se parecen algunos nucleótidos de la primera mitad de este elemento promotor (2 y 4), siendo el resto de la secuencia diferente; aunque en las posiciones 8, 9 y 11 puede observarse la preferencia por una purina (G o A) tanto en los parásitos como en *Xenopus*. Como se mencionó en antecedentes, se han identificado previamente las secuencias de

las cajas promotoras del RNAr 5S en tripanosomátidos (Hernández-Rivas *et al*, 1992, Cordingley, 1985), aunque estas regiones no han sido caracterizadas funcionalmente. Al igual que en el modelo de *Xenopus*, un factor homólogo a TFIIIA en tripanosomátidos podría estarse uniendo a las cajas A, C y el EI durante la transcripción del RNAr 5S. Cabe mencionar que muchos de los factores de transcripción de Pol III en tripanosomátidos aún no se han identificado, incluido TFIIIA (Das *et al*, 2008). En general, los tres elementos se encuentran conservados entre tripanosomátidos, por lo que el factor de unión a estas regiones podría estar también altamente conservado entre estos organismos.

Aunque la mayoría de los genes de RNAr 5S de eucariontes contienen las cajas A y C y el EI, muchos contienen secuencias adicionales dentro del gen y río arriba del mismo que afectan significativamente su transcripción, por ejemplo en levadura, *Acanthamoeba*, *Drosophila*, etc. (Peng y Bateman, 2004). En este sentido, la perspectiva es amplia en tripanosomátidos para realizar experimentos que permitan determinar si existen regiones adicionales a los promotores clásicos.

Se identificaron también las estructuras secundarias de los RNAr 5S en tripanosomátidos y sus elementos distintivos. En otros eucariontes se ha observado que la hélice I es esencial para la estabilización de la molécula y su integración al ribosoma, así como para su interacción con las proteínas L5 y TFIIIA (Fig. 27D). Por su parte, el asa B – hélice III y el asa C son cruciales para la interacción con proteínas ribosomales L5 y L18, así como para el contacto entre subunidades ribosomales. Las hélices IV y el asa D son importantes debido a su interacción con el RNAr 28S y participan en la integración de la subunidad mayor ribosomal. Por su parte, el asa E interactúa con factores de traducción y transcripción (TFIIIA) y también se une al RNAr 28S (Smirnov *et al*, 2008). Estas estructuras fueron localizadas en los RNAr 5S de los tripanosomátidos y podrían ejercer funciones muy similares (Fig. 27). Al igual que en otros eucariontes, en tripanosomátidos el RNAr 5S podría funcionar a manera de enlace entre diferentes regiones del ribosoma e interactuar con distintas proteínas y factores de transcripción y traducción.

7. CONCLUSIONES

- El número de genes de RNAt en tripanosomátidos es de 83, 66 y 120 para *L. major*, *T. brucei* y *T. cruzi*, respectivamente. Estos números varían ligeramente de lo anotado en las bases de datos (82, 65 y 115, respectivamente).
- En la base de datos de *T. brucei* existen genes mal anotados: Tb08_tRNA_Leu_2 (Tb927.8.6564) y Tb08_tRNA_Gln_2 (Tb927.8.6566) se encuentran anotados en la cadena sentido y en realidad se encuentran en la antisentido. El gen anotado como Sec (Tb04_tRNA_Sec_1) se trata en realidad del sRNA76.
- Se localizaron 1, 2 y 8 copias del RNAt-Sec en los genomas de *L. major*, *T. brucei* y *T. cruzi*, respectivamente.
- En *T. cruzi* se encontraron 4 secuencias de RNAt con isotipo indeterminado que se eliminaron de la lista de genes de RNAt.
- Se encontraron 46 RNAt isoaceptores en tripanosomátidos, 45 correspondientes a los aminoácidos comunes y 1 correspondiente a Sec. Números similares de isoaceptores se encuentran en eucariontes como humano, chimpancé y *C. elegans*.
- Los Trypanosoma utilizan la estrategia de recambio A1 – G1 para leer los codones cuyo RNAt correspondiente no se encuentra en el genoma.
- Se localizaron los genes de RNAt-Met iniciador y elongador y se identificó la mayoría de las características de secuencia propias de ambos tipos de RNAt-Met.
- Tres pares de G:C consecutivas en el tallo del anticodón se encuentran típicamente en el RNAt-i en eucariontes; sin embargo, en tripanosomátidos se encontraron en ambos tipos de RNAt-Met, por lo que estos parásitos podrían utilizar secuencias diferentes para distinguir el RNAt-i del RNAt-e.
- En *L. major* y *T. brucei*, la mayoría de los genes de RNAt están organizados en clusters de 2-10 genes, los cuales pueden contener otros genes transcritos por Pol III como los RNAsn o, en el caso de *L. major*, RNAr 5S.
- En general, las regiones que separan los genes de RNAt son cortas (327 pb en *T. brucei* y 202 pb en *L. major*, en promedio).

- No se encontró sintenia en la mayoría de los clusters de genes de RNAt entre las tres especies de tripanosomátidos analizadas, a pesar de que los genes codificantes de proteínas son altamente sinténicos.
- Sólo los genes de RNAt-Tyr presentaron un intrón de 11 (*L. major*, *T. brucei*) ó 13 pb (*T. cruzi*).
- Reportamos la secuencia consenso para los elementos promotores de los RNAt en tripanosomátidos: cajas A y B. Al separar los genes en clase I y clase II se observaron secuencias consenso más conservadas.
- Reportamos la secuencia consenso de los probables elementos promotores internos de los RNAt-Sec: la caja A presenta una inserción y la caja B presenta dos cambios, con respecto a los RNAt clásicos.
- Como señal de término de la transcripción de los RNAt, *L. major* y *T. brucei* poseen en promedio una secuencia de de 5 Ts (± 1), mientras que *T. cruzi* contiene en promedio 7 Ts (± 2).
- En *T. cruzi* el 48% de los genes de RNAt presenta una segunda región de Ts, mientras que en *L. major* y *T. brucei* la presentan sólo el 16% y 27% de los genes de RNAt, respectivamente.
- Se encontró 1 gen de RNAt isodecodificador para *L. major*, 2 para *T. brucei* y 6 para *T. cruzi*. Variaciones en las secuencias se encuentran dentro o cerca de los promotores internos, por lo que podría existir una regulación diferencial de la transcripción por Pol III.
- Se obtuvo una correlación positiva entre el número de genes de RNAt y el uso de codones en las tres especies de tripanosomátidos.
- En las bases de datos se encuentran anotados 11, 9 y 135 genes de RNAr 5S en *L. major*, *T. brucei* y *T. cruzi*, respectivamente.
- Se obtuvieron secuencias consenso de los elementos promotores internos (cajas A, C y el EI) del RNAr 5S en tripanosomátidos. En general, estos elementos se encuentran bastante conservados entre tripanosomátidos.

8. BIBLIOGRAFIA

- Agris, P., Vendeix, F. y Graham, W. 2007. tRNA's Wobble decoding of the Genome: 40 years of modification. *Journal of Molecular Biology* 366 (1): 1-13.
- Agris, P. 2004. Decoding the genome: a modified view. *Nucleic Acids Research*. 32 (1): 223-238.
- Allmang, C. y Krol, A. 2006. Selenoprotein synthesis: UGA does not end the story. *Biochimie* 88 (11): 1561-1571
- Aström S., von Pawel-Rammingen, U. y Byström A. 1993. The yeast initiator tRNA Met can act as an elongator tRNA (Met) *in vivo*. *Journal of Molecular Biology* 233(1):43-58.
- Barciszewska, M., Szymanski, M., Erdmann, V. y Barciszewski. J. 2000. Ribosomal RNA 5S. *Biomacromolecules* 1 (3): 297-302.
- Bates, P. 2007. Transmission of *Leishmania* metacyclic promastigotes by phlebotomine sand flies. *International Journal for Parasitology* 37 (10): 1097-1106.
- Béjà, O., Ullu, E. y Michaeli, S. 1993. Identification of a tRNA-like molecule that copurifies with the 7SL RNA of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 57 (2): 223-230.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Böhme, U., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C., Arrowsmith, C., Atkin, R. J., Barron, A. J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T. J., Churcher, C., Clark, L. N., Corton, C. H., Cronin, A., Davies, R. M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M. C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B. R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A. X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P. J., Moule, S., Martin, D. M., Morgan, G. W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C. S., Peterson, J., Quail, M. A., Rabinowitsch, E., Rajandream, M. A., Reitter, C., Salzberg, S.L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A. J., Tallon, L., Turner, C. M., Tait, A., Tivey, A. R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M. D., Embley, T. M., Gull, K., Ullu, E., Barry, J. D., Fairlamb, A. H., Opperdoes, F., Barrell, B. G., Donelson, J. E., Hall, N., Fraser, C. M., Melville, S. E., El-Sayed, N. 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309 (5733): 416-422
- Braglia, P., Percudani, R. y Dieci, G. 2005. Sequence context effects on Oligo(dT) termination signal recognition by *Saccharomyces cerevisiae* RNA polymerase III. *Journal of Biological Chemistry* 280 (20): 19551-19562.
- Brennan, T. y Sundaralingam, M. 1976. Structure of transfer RNA molecules containing the long variable loop. *Nucleic Acids Research* 3 (11): 3235-3251.
- Burleigh y Woolsey. 2002. Cell signaling and *T. cruzi* invasion. *Cellular Microbiology* 4 (11): 701-711.
- Campbell, D., Thomas, S. y Sturm, N. 2003. Transcription in kinetoplastid protozoa: why be normal?. *Microbes and Infection* 5 (13): 1231-1240.
- Cassago, A., Rodrigues, E.M., Prieto, E.L., Gaston, K.W., Alfonso, J.D., Iribar, M.P., Berry, M.J., Cruz, A.K. y Thiemann, O.H. 2006. Identification of *Leishmania* selenoproteins and SECIS element. *Molecular and Biochemical Parasitology* 149 (2): 128-134.
- Carbon, P. y Krol, A. 1991. Transcription of the *Xenopus laevis* selenocysteine tRNA (Ser)^{Sec} gene: a system that combines an internal B box and upstream elements also found in U6 snRNA genes. *The EMBO Journal* 10 (3): 599-606.
- Carlton, J., Silva, J. y Hall, N. 2005. The genome of model malaria parasites, and comparative genomics. *Current Issues in Molecular Biology* 7 (1): 23-27

Cognat, V., Deragon, J.M., Vinogradova, E., Salinas, T., Remacle, C. y Maréchal-Drouard, L. 2008. On the evolution and expression of *Chlamydomonas reinhardtii* nucleus-encoded transfer RNA genes. *Genetics* 179 (1): 113-123.

Cordingley, J.S. 1985. Nucleotide sequence of the 5S ribosomal RNA gene repeat of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 17(3): 321-330.

Das, A., Banday, M y Bellofatto, V. 2008. RNA polymerase Transcription Machinery in Trypanosomes. *Eukaryotic Cell* 7 (3): 429-434.

Desjeux, P. 2004. Leishmaniasis: current situation and new perspectives. *Comparative Immunology, Microbiology and Infectious Diseases* 27 (5): 305-318.

Dieci, G., Fiorino, G., Castelnovo, M., Teichmann, M. y Pagano, A. 2007. The expanding RNA polymerase III transcriptome. *Trends in Genetics* 23 (12): 614-622.

Dinman, J. 2005. 5S rRNA: Structure and Function from Head to Toe. *International Journal of Biomedical Sciences* 1(1): 2-7.

Dittmar, K., Goodenbour, J. y Pan, T. 2006. Tissue-specific differences in human transfer RNA expression. *PLoS Genetics* 2 (12): 2107-2115.

Drabkin, H. Estrella, M. y Rajbhandary, U. 1998. Initiator-Elongator Discrimination in vertebrate tRNAs for Protein Synthesis. *Molecular and Cellular Biology* 18 (3): 1459-1466.

Drabkin H.J., Helk, B. y Rajbhandary, U.L. 1993. The role of nucleotides conserved in eukaryotic initiator methionine tRNAs in initiation of protein synthesis. *Journal of Biological Chemistry* 268(33):25221-8.

Duret, L., 2000. tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. *Trends in Genetics* 16 (7): 287-289.

El-Sayed, N. Myler, P., Bartholomeu, D., Nilsson, D., Aggarwal, G., Tran, A., Ghedin, E., Worthey, E. A., Delcher, A.L., Blandin, G., Westenberger, S., Caler, E., Cerqueira, G.C., Branche, C. Haas, B., Anupama, A., Arner, E., Åslund, E., Attipoe, P., Bontempi, E., Bringaud, F., Burton, P., Cadag, P., Campbell, D., Carrington, M., Crabtree, J. Darban, J., da Silveira, F., de Jong, P., Edwards, K., Englund, P.T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A.C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M.J., Lorenzi, H., Louie, T., Machado, C.R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J.L., Rinta, J., Robertson, L., Salzberg, S.L., Sanchez, D.O., Seyler, A., Sharma, R., Shetty, J., Simpson, A.J., Sisk, A., Tammi, M.T., Tarleton, R., Teixeira, S., Aken, S. V., Vogt, C., Ward, P.N., Wickstead, B., Wortman, J., White, O., Fraser, C.M., Stuart, K.D. y Andersson, B. 2005. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309 (5733): 409-415.

El-Sayed, N., Myler, P., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, C. Haas, B., Tran, A., Wortman, J., Alsmark, U.C., Angiuoli, S., Anupama, A., Badger, J., Bringaud, F., Cadag, E., Carlton, J., Cerqueira, Creasy, T., Delcher, A., Djikeng, A., Embley, T.M., Hauser, C., Ivens, A., Kummerfeld, S., Pereira-Leal, J., Nilsson, D., Peterson, J., Salzberg, S., Shallom, J., Silva, J., Sundaram, J., Westenberger, S., White, O., Melville, S., Donelson, J., Andersson, B., Stuart, K. y Hall, N. 2005. Comparative Genomics of Trypanosomatid Parasitic Protozoa. *Science* 309 (5733): 404-409.

Farruggio D., Chaudhuri J., Maitra U. y Rajbhandary U. 1996. The A1 x U72 base pair conserved in eukaryotic initiator tRNAs is important specifically for binding to the eukaryotic translation initiation factor eIF2. *Molecular Cell Biology* 16(8):4248-56.

Goodenbour, J. y Pan, T. 2006. Diversity of tRNA genes in eukaryotes. *Nucleic Acids Research* 34(21): 6137-6146.

Haeusler, R. y Engelke, D. 2006. Spatial organization of transcription by RNA polymerase III. *Nucleic Acids Research* 34 (17): 4826-4836.

Hasan, G., Turner, M. y Cordingley, J. 1984. Ribosomal RNA genes of *Trypanosoma brucei*: Mapping the regions specifying the six small ribosomal RNAs. *Gene* 27 (1): 75-86

Hatfield, D. y Gladyshev, V. 2002. How selenium has altered our understanding of the genetic code. *Molecular and Cellular Biology* 22 (11): 3565-3576.

Hellman, K., Ciganda, M., Brown, S., Li, J., Ruyechan, W. y Williams, N. 2007. Two trypanosome-specific proteins are essential factors for 5S rRNA abundance and ribosomal assembly in *Trypanosoma brucei*. *Eukaryotic Cell* 6 (10): 1766-1772.

Hernández-Rivas, R., Florencio-Martínez, L. Martínez-Salazar, M. y Martínez-Calvillo, S. 2007. Gene expresión and transcriptional machinery in trypanosomatid and apicomplexa parasites. *Advances in the Immunobiology of Parasitic Diseases*: 313:337.

Hernández-Rivas, R., Martínez-Calvillo, S., Romero, M. y Hernández, R. 1992. *Trypanosoma cruzi* 5S rRNA genes: molecular cloning, structure and chromosomal organization. *FEMS Microbiology Letters* 92 (1): 63-68.

Horn, D. 2008. Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids. *BMC Genomics* 9 (2).

Hubert, N., Walczak, R., Sturchler, C., Myslinkski, E., Schuster, C., Westhof, E., Carbon, P. y Krol, A. 1996. RNAs mediating cotranslational insertion of selenocysteine in eukaryotic selenoproteins. *Biochimie* 78 (7): 590-596

Ikemura, T. 1985. Codon usage and tRNA content in Unicellular and Multicellular Organisms. *Molecular Biology and Evolution*. 2 (1): 13-34.

Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M. A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S.M., Bianchettin, G., Borzym, K., Bothe, G., Bruschi, C.V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R. M., Cronin, A., Cruz, A. K., Davies, R. M., De Gaudenzi, J., Dobson, D. E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A. C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J. C., Müller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'neil, S., Pentony, M., Pohl, T. M., Price, C., Purnelle, B., Quail, M. A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J. C., Rutter, S., Saunders, D., Schäfer, M., Schein, J., Schwartz, D. C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D. F., Blackwell, J. M., Stuart, K. D., Barrell, B. y Myler, P. J. 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309 (5733): 436-42.

Johansson, L., Gafvelin, G. Arnér, E. 2005. Selenocysteine in proteins – properties and biotechnological use. *Biochimica et Biophysica Acta* 1726 (1): 1-13.

Kanaya, S., Yamada, Y., Kinouchi, M., Kudo, Y. e Ikemura, T. 2001. Codon usage and tRNA genes in eukaryotes: correlation of codon usage diversity with translation efficiency and with CG-dinucleotide usage as assessed by multivariate analysis. *Journal of Molecular Evolution* 53 (4-5): 290-298.

Kanaya, S., Yamada, Y., Kudo, Y. e Ikemura, T. 1999. Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of *Bacillus subtilis* tRNAs: gene expression level and species-specific diversity of codon usage based on multivariate analysis. *Gene* 238 (1): 145-155.

Kapp, L, Kolitz, S. y Lorsch, J. 2006. Yeast initiator tRNA identity elements cooperate to influence multiple steps of translation initiation. *RNA* 12 (5): 751-764.

- Kryukov, G., Castellano, S., Novoselov, S., Lobanov, A., Zehtab, O., Guigo, G. y Gladyshev, V. 2003. Characterization of Mammalian Selenoproteomes. *Science* 300 (5624): 1439-1443.
- Lee, B., Kang, S. y Hatfield, D. 1989. Transcription of *Xenopus* selenocystein tRNA-Ser gene is directed by multiple 5'-extragenic regulatory elements. *Journal of Biological Chemistry* 264 (16): 9696-9702.
- Lowe, T. y Eddy, S. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* 25 (5): 955-964.
- Liu, L, Ben-Shlomo, H., Xu, Y., Zeev Stern, Goncharov, I., Zhang, Y. y Michaeli, S. 2003. The Trypanosomatid Signal Recognition Particle Consists of Two RNA Molecules, a 7SL RNA Homologue and a Novel tRNA-like Molecule. *Journal of Biological Chemistry* 278 (20): 18271-18280.
- Mandal, N., Mangroo, D., Dalluge, J., McCloskey, J. y Rajbhandary, U. 1996. Role of three consecutive G:C base pairs conserved in the anticodon stem of initiator tRNAs in initiation of protein synthesis in *Escherichia coli*. *RNA* 2 (5): 473-482.
- Mans, K., Pleij, C. y Bosch, L. 1991. tRNA-like structures. Structure, function and evolutionary significance *European Journal of Biochemistry* 201 (2): 303-324
- Marck, C. y Grosjean, H. 2002. tRNomics: Analysis of tRNA genes from 50 genomes of Eukarya, Archea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. 8 (10) : 1189-1232.
- Martínez-Calvillo, S., Nguyen, D., Stuart, K. y Myler, P. 2004. Transcription initiation and termination on *Leishmania major* chromosome 3. *Eukaryotic Cell* 3 (2):506-517.
- Martínez-Calvillo, S., Sunkin, S., Yan, S., Fox, M., Stuart, K. y Myler, P. 2001. Genomic organization and functional characterization of the *Leishmania major* Friedlin ribosomal RNA gene locus. *Molecular and Biochemical Parasitology* 116 (2): 147-157.
- Murray, H., Berman, J., Davies, C. y Saravia, N. 2005. Advances in leishmaniasis. *Lancet* 366 (9496): 1561-1577.
- Morrison, H., McArthur, A., Gillin, F., Aley, S.B., Adam, R.D., Olsen, G.J., Best, A.A. Zacheus-Cande, W., Chen, F., Cipriano, M. J., Davids, B., Dawson, S.C., Elmendorf, H., Hehl, A.B., Holder, M. E., Huse, S. M., Kim, U. U., Lasek-Nesselquist, E., Manning, G., Nigam, A., Nixon, J.E.J., Palm, D., Passamaneck, N. E., Prabhu, A., Reich, C., Reiner, D. S., Samuelson, D., Svard, S. y Sogin, M.. 2007. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. *Science* 317 (5846): 1921-1926.
- Mottram, J., Bell, S., Nelson, R. y Barry, J. D. 1991. tRNAs of *Trypanosoma brucei*. *Journal of Biological Chemistry* 266 (27): 18313-18317
- Myslinski E., Krol, A., Carbon, P. 1991. Optimal tRNA Ser-Sec gene activity requires an upstream SPH motif. *Nucleic Acids Research* 20 (2):203-209.
- Myslinski, E., Schuster, C., Huet, J., Sentenac, A., Krol, A. y Carbon, P. 1993. Point mutations 5' to the tRNA selenocysteine TATA box alter RNA pol III transcriptiion by affecting the binding of TBP. *Nucleic Acids Research* 21 (25): 5852-5858.
- Nakaar, V., Dare, A., Hong, D., Ullu, E. y Tshudi, C. 1994. Upstream tRNA genes are essential for expression of small nuclear and cytoplasmic RNA genes in trypanosomes. *Molecular and cellular biology* 14 (10): 6736-6742.
- Obado, S., Bot, C., Nilsson, D., Andersson, B. y Kelly, J. M. 2007. Repetitive DNA is associated with centromeric domains in *Trypanosoma brucei* but not *Trypanosoma cruzi*. *Genome Biology* 8 (3): R37.
- Park, J., Choi, I., Kang, S., Lee, J., Hatfield, D. y Lee, B. 1995. Upstream promoter elements are sufficient for selenocysteine tRNA Ser Sec gene transcription and to determine the transcription start point. *Gene* 162 (1): 13-19.

- Park, J., Yang, E., Hatfield, D. y Lee, B. 1996. Analysis of the Selenocysteine tRNA Ser Sec gene transcription in vitro using *Xenopus* oocyte extracts. *Biochemical and Biophysical Research Communications*. 226 (1): 231-236.
- Paule, M. y White, R. 2000. Transcription by RNA polymerase I and III. *Nucleic Acids Research* 28 (6): 1283-1298.
- von Pawel-Rammingen U., Aström S. y Byström A.S. 1992. Mutational analysis of conserved positions potentially important for initiator tRNA function in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*. 12(4):1432-42
- Peng, Z. y Bateman, E. 2004. Analysis of the 5S rRNA gene promoter from *Acanthamoeba castellanii*. *Molecular Microbiology* 52 (4): 1123-1132.
- Percudani, R. 2001. Restricted wobble rules for eukaryotic genomes. *Trends in Genetics* 17(3):133-5.
- Percudani, R., Pavesa, A. y Ottonello, S. 1997. Transfer RNA gene redundancy and translational selection in *Saccharomyces cerevisiae*. *Journal of Molecular Biology* 268 (2): 322-330.
- Prescott, E. y Proudfoot, N. 2002. Transcriptional collision between convergent genes in budding yeast. *PNAS* 99 (13): 8796-8801.
- Rocha, E. 2004. Codon usage bias from tRNA's point of view: redundancy, specialization and efficient decoding for translation optimization. *Genome research* 14 (11): 2279-2286.
- Schneider, A., McNally, P. y Agabian, N. 1993. Splicing and 3'-processing of the Tyrosine tRNA of *Trypanosoma brucei*. *Journal of Biological Chemistry* 268 (29): 21868-21874.
- Schramm, L. y Hernandez, N. 2002. Recruitment of RNA polymerase III to its target promoters. *Genes and Development* 16 (20): 2593-2620.
- Schuster, C., Myslinski, E., Krol, A. y Carbon, P. 1995. Staf, a novel zinc finger protein that activates the RNA polymerase III promoter of the selenocysteine tRNA gene. *The EMBO Journal* 14 (15): 3777-3787.
- Scott, K., White, C. y Huntington, W. 2007. An RNA Polymerase III-Dependent Heterochromatin Barrier at Fission Yeast Centromere 1. *PLoS ONE* 2(10):e1099.
- Smirnov, A., Entelis, N., Krasheninnikov, I., Martin, R. y Tarassov, I. 2008. Specific features of 5S rRNA structure – its interactions with macromolecules and possible functions. *Biochemistry* 73 (13): 1418-1437.
- Stortchevoi A., Varshney U. y RajBhandary, U. 2003. Common location of determinants in initiator transfer RNAs for initiator-elongator discrimination in bacteria and in eukaryotes. *Journal of Biological Chemistry* 278(20):17672-9.
- Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gürtler, R., McKerrow, J., Reed, S. y Tarleton, R. 2008. Kinetoplastids: related protozoan pathogens, different diseases. *The Journal of Clinical Investigation* 118 (4): 1301-1310.
- Szymanski, M., Barciszewska, M., Erdmann, V. y Barciszewski, J. 2003. 5S rRNA: structure and interactions. *Biochemical Journal* 371 (Pt 3): 641- 651.
- Tan, T., Pach, R., Crausaz, A., Ivenz, A. y Scheneider, A. 2002. tRNAs in *Trypanosoma brucei*: Genomic organization, expression and mitochondrial import. *Molecular and Cellular Biology* 22 (11): 3707-16.
- Thompson, M., Haeusler, R., Good, P. y Engelke, D. 2003. Nucleolar clustering of dispersed tRNA genes. *Science* 302 (5649): 1399-1401.
- Tripathy, S. y Tyler, B.M. 2006. The repertoire of transfer RNA is tuned to codon usage bias in the genomes of *Phytophthora sojae* and *Phytophthora ramorum*. *Molecular Plant - Microbe Interactions* 19 (12): 1322-1328.
- Wickstead, B., Ersfeld, K. y Gull, K. 2003. Repetitive Elements in Genomes of Parasitic Protozoa. *Microbiology and Molecular Biology Reviews*, 67 (3): 360-375.

9. APENDICES

APÉNDICE I: Datos completos sobre la localización de los genes de RNAt en Tripanosomátidos (Complementaria a Fig. 13 y 14)

Leishmania major

Gen	Anticodón	Coordenadas	Cadena de DNA
LmjF03.TRNALYS.01	CTT	259317..259389	Sentido
LmjF05.TRNAARG.01	ACG	359954..360025	Sentido
LmjF07.TRNAARG.01	ACG	381407..381478	Antisentido
LmjF09.TRNAARG.01	CCG	418949..419020	Sentido
LmjF09.TRNAGLU.01	TTC	418491..418562	Antisentido
LmjF09.TRNAHIS.01	GTG	405864..405934	Sentido
LmjF09.TRNAHIS.02	GTG	419087..419157	Sentido
LmjF09.TRNALEU.01	CAG	277210..277292	Sentido
LmjF09.TRNAMET.01	CAT	277080..277151	Sentido
LmjF09.TRNAPHE.01	GAA	406115..406187	Antisentido
LmjF09.TRNAVAL.01	CAC	405537..405609	Sentido
LmjF09.TRNAVAL.02	CAC	418802..418874	Antisentido
LmjF10.TRNAASN.01	GTT	533324..533396	Antisentido
LmjF10.TRNAGLY.01	GCC	533197..533267	Antisentido
LmjF10.TRNALYS.01	CTT	533082..533154	Sentido
LmjF11.TRNAALA.01	CGC	157119..157190	Sentido
LmjF11.TRNAALA.02	CGC	162729..162800	Antisentido
LmjF11.TRNAARG.01	ACG	381490..381561	Sentido
LmjF11.TRNAGLY.01	CCC	157252..157322	Sentido
LmjF11.TRNALEU.01	AAG	156964..157045	Antisentido
LmjF11.TRNALEU.02	AAG	162874..162955	Sentido
LmjF11.TRNAMET.01	CAT	381341..381413	Antisentido
LmjF15.TRNAASP.01	GTC	322953..323024	Sentido
LmjF15.TRNAGLU.01	CTC	323065..323136	Antisentido
LmjF16.TRNAGLN.01	CTG	455436..455507	Antisentido
LmjF17.TRNAALA.01	AGC	344275..344347	Sentido
LmjF17.TRNAASP.01	GTC	343984..344055	Antisentido
LmjF17.TRNASER.01	GCT	344133..344216	Sentido
LmjF21.TRNALYS.02	CTT	449408..449480	Sentido
LmjF21.TRNAPRO.01	AGG	449049..449120	Sentido
LmjF21.TRNASER.01	GCT	166347..166427	Antisentido
LmjF21.TRNAVAL.01	AAC	449272..449345	Sentido
LmjF23.TRNAARG.01	ACG	225746..225817	Antisentido
LmjF23.TRNAARG.02	TCG	226329..226400	Sentido
LmjF23.TRNAGLN.01	TTG	225427..225498	Antisentido
LmjF23.TRNAGLY.01	TCC	226826..226897	Antisentido
LmjF23.TRNAILE.01	TAT	225169..225242	Sentido
LmjF23.TRNALEU.01	CAG	226042..226124	Antisentido
LmjF23.TRNALYS.01	TTT	226715..226787	Sentido

LmjF23.TRNATHR.01	TGT	226185..226256	Antisentido
LmjF23.TRNATRP.01	CCA	226973..227044	Sentido
LmjF23.TRNAVAL.01	TAC	225287..225359	Antisentido
LmjF24.TRNAASP.01	GTC	680991..681062	Antisentido
LmjF24.TRNAGLN.01	CTG	621500..621571	Sentido
LmjF24.TRNAILE.01	AAT	621952..622025	Sentido
LmjF24.TRNAILE.02	AAT	623347..623420	Sentido
LmjF24.TRNALEU.01	TAA	621785..621866	Antisentido
LmjF24.TRNAPRO.01	CGG	205318..205389	Sentido
LmjF29.TRNALEU.01	CAA	819678..819759	Antisentido
LmjF29.TRNASER.01	CGA	819818..819898	Antisentido
LmjF30.TRNATHR.01	CGT	785495..785566	Antisentido
LmjF31.TRNAALA.01	AGC	203939..204011	Sentido
LmjF31.TRNAGLU.01	CTC	494758..494829	Antisentido
LmjF31.TRNAGLY.01	GCC	494642..494712	Sentido
LmjF31.TRNAPHE.01	GAA	1315562..1315634	Sentido
LmjF31.TRNASER.01	AGA	1315687..1315767	Antisentido
LmjF33.TRNAALA.01	TGC	106939..107010	Sentido
LmjF33.TRNAARG.01	CCT	106658..106730	Antisentido
LmjF33.TRNAARG.02	TCT	106790..106862	Antisentido
LmjF34.TRNAASN.01	GTT	1534453..1534525	Antisentido
LmjF34.TRNAASN.02	GTT	1535020..1535092	Sentido
LmjF34.TRNAILE.01	AAT	470433..470506	Sentido
LmjF34.TRNALEU.01	TAG	470676..470755	Antisentido
LmjF34.TRNAMET.01	CAT	470568..470640	Sentido
LmjF34.TRNAPRO.01	CGG	1534885..1534956	Sentido
LmjF34.TRNASER.01	TGA	470830..470910	Sentido
LmjF34.TRNATHR.01	CGT	1534732..1534803	Antisentido
LmjF34.TRNATYR.01	GTA	1534584..1534667	Antisentido
LmjF34.TRNAVAL.01	AAC	470126..470199	Antisentido
LmjF36.TRNACYS.01	GCA	2469004..2469075	Sentido
LmjF36.TRNAGLN.01	CTG	1607663..1607734	Antisentido
LmjF36.TRNAGLY.01	GCC	1032519..1032589	Sentido
LmjF36.TRNAGLY.02	GCC	1033367..1033437	Antisentido
LmjF36.TRNALEU.01	AAG	2469866..2469947	Antisentido
LmjF36.TRNAMET.01	CAT	2470010..2470081	Antisentido
LmjF36.TRNAPRO.01	AGG	1033511..1033582	Sentido
LmjF36.TRNAPRO.02	TGG	2470154..2470225	Sentido
LmjF36.TRNATHR.01	AGT	491737..491808	Antisentido
LmjF36.TRNATHR.02	AGT	492633..492704	Sentido
LmjF36.TRNATHR.03	AGT	1033233..1033304	Antisentido
LmjF36.TRNATYR.01	GTA	491573..491656	Antisentido
LmjF36.TRNATYR.02	GTA	492788..492871	Sentido
Chr06.Sec	TCA	69586..69673	Antisentido

Trypanosoma brucei

Gen	Anticodón	Coordenadas	Cadena de DNA
Tb04_tRNA_Glu_1 Tb927.4.3292	CTC	860615..860686	Antisentido
Tb04_tRNA_Leu_1 Tb927.4.3196	TAG	842789..842868	Antisentido
Tb04_tRNA_Met_1 Tb927.4.3199	CAT	842927..842999	Antisentido
Tb04_tRNA_Met_2 Tb927.4.3294	CAT	860774..860846	Sentido
Tb04_tRNA_Pro_1 Tb927.4.1195	CGG	317914..317985	Antisentido
Tb04_tRNA_Ser_1 Tb927.4.3296	TGA	860890..860970	Antisentido
Tb04_tRNA_Thr_1 Tb927.4.1216	CGT	325958..326029	Sentido
Tb04_tRNA_Tyr_1 Tb927.4.1219	GTA	326098..326181	Antisentido
Tb04_tRNA_Val_1 Tb927.4.3193	AAC	842652..842725	Antisentido
Tb04_tRNA_Val_2 Tb927.4.3298	AAC	861087..861160	Sentido
Tb04_tRNA_Val_3 Tb927.4.4666	CAC	1287445..1287517	Antisentido
Tb05_tRNA_Arg_1 Tb927.5.445	ACG	138373..138444	Antisentido
Tb05_tRNA_Glu_1 Tb927.5.2765	CTC	874266..874337	Sentido
Tb07_tRNA_Ala_1 Tb927.7.6821	AGC	1890511..1890583	Antisentido
Tb07_tRNA_Ala_2 Tb927.7.6826	AGC	1894352..1894424	Sentido
Tb07_tRNA_Ala_3 Tb927.7.7175	CGC	2059249..2059320	Antisentido
Tb07_tRNA_Asp_1 Tb927.7.6823	GTC	1890810..1890881	Sentido
Tb07_tRNA_Asp_2 Tb927.7.6824	GTC	1894054..1894125	Antisentido
Tb07_tRNA_Gly_1 Tb927.7.1175	GCC	300549..300619	Sentido
Tb07_tRNA_Lys_1 Tb927.7.6845	CTT	1910941..1911013	Sentido
Tb07_tRNA_Ser_1 Tb927.7.6822	GCT	1890653..1890733	Antisentido
Tb07_tRNA_Ser_2 Tb927.7.6825	GCT	1894202..1894282	Sentido
Tb08_tRNA_Arg_1 Tb927.8.2859	TCG	858599..858670	Sentido
Tb08_tRNA_Arg_2 Tb927.8.2863	ACG	861528..861599	Antisentido
Tb08_tRNA_Arg_3 Tb927.8.5345	ACG	1587504..1587575	Sentido
Tb08_tRNA_Asn_1 Tb927.8.2663	GTT	793455..793527	Sentido
Tb08_tRNA_Cys_1 Tb927.8.3523	GCA	1053326..1053397	Antisentido
Tb08_tRNA_Cys_2 Tb927.8.3526	GCA	1053540..1053611	Sentido
Tb08_tRNA_Gln_1	TTG	854457..854528	Sentido

Tb927.8.2852			
Tb08_tRNA_Gln_2 Tb927.8.6566	CTG	1892288..1892359	Sentido
Tb08_tRNA_Gln_3 Tb927.8.6568	CTG	1892449..1892520	Sentido
Tb08_tRNA_Gly_1 Tb927.8.2856	TCC	858106..858177	Sentido
Tb08_tRNA_Ile_1 Tb927.8.2851	TAT	854289..854362	Antisentido
Tb08_tRNA_Ile_2 Tb927.8.6562	AAT	1891988..1892061	Sentido
Tb08_tRNA_Ile_3 Tb927.8.6572	AAT	1892600..1892673	Sentido
Tb08_tRNA_Leu_1 Tb927.8.2857	CAG	858265..858346	Sentido
Tb08_tRNA_Leu_2 Tb927.8.6564	TAA	1892111..1892192	Sentido
Tb08_tRNA_Lys_1 Tb927.8.2854	CTT	854728..854800	Sentido
Tb08_tRNA_Lys_2 Tb927.8.2862	CTT	861366..861438	Sentido
Tb08_tRNA_Thr_1 Tb927.8.2858	TGT	858430..858501	Antisentido
Tb08_tRNA_Trp_1 Tb927.8.2666	CCA	793613..793684	Antisentido
Tb08_tRNA_Val_1 Tb927.8.2853	TAC	854585..854657	Sentido
Tb09_tRNA_Gly_1	GCC	2036843..2036913	Sentido
Tb10_tRNA_Ala_1	TGC	2603567..2603638	Antisentido
Tb10_tRNA_Arg_1	TCT	2603741..2603813	Sentido
Tb10_tRNA_Arg_2	CCT	2604030..2604102	Sentido
Tb10_tRNA_Asn_1	GTT	2603416..2603488	Sentido
Tb10_tRNA_Gly_1	GCC	1867515..1867585	Sentido
Tb10_tRNA_Leu_1	AAG	1907650..1907731	Sentido
Tb10_tRNA_Lys_1	TTT	2603873..2603944	Sentido
Tb10_tRNA_Met_1	CAT	1907793..1907864	Antisentido
Tb10_tRNA_Pro_1	TGG	1907937..1908008	Sentido
Tb10_tRNA_Pro_2	AGG	1904001..1904072	Sentido
Tb10_tRNA_Thr_1	AGT	1904133..1904204	Sentido
Tb11_tRNA_Ala_1	CGC	3469847..3469918	Antisentido
Tb11_tRNA_Arg_1	CCG	3469985..3470056	Antisentido
Tb11_tRNA_Glu_1	TTC	3452226..3452297	Sentido
Tb11_tRNA_Gly_1	CCC	1874612..1874682	Sentido
Tb11_tRNA_His_1	GTG	3452094..3452164	Sentido
Tb11_tRNA_Phe_1	GAA	3452357..3452429	Sentido
Tb11_tRNA_Phe_2	GAA	3470168..3470240	Sentido
Tb927.3.4393 (chr3-Ser)	AGA	1233524..1233604	Sentido
Tb927.3.4396 (chr3-Leu)	CAA	1233810..1233891	Antisentido
Tb927.3.4399 (chr3-Ser)	CGA	1233951..1234031	Antisentido

Chr09.Sec.1	TCA	505606-505693	Antisentido
Chr09.Sec.2	TCA	509443-509530	Antisentido

Trypanosoma cruzi

Gen	Aminoácido	Anticodón	Contig	Coordenadas	Cadena de DNA
Tc00.1047053510057.40	ALA	AGC	5833	7203..7275	Antisentido
Tc00.1047053508909.130	ALA	AGC	7924	35863..35935	Sentido
Tc00.1047053506599.3	ALA	CGC	7016	1187..1261	Antisentido
Tc00.1047053508873.408	ALA	CGC	7909	90692..90766	Sentido
Tc00.1047053475029.40	ALA	TGC	4041	754..825	Sentido
Sin nombre (no anotado)	ALA	TGC	8001	28744..28815	Antisentido
Tc00.1047053504427.243	ARG	ACG	6288	53948..54019	Antisentido
Tc00.1047053506619.59	ARG	ACG	7024	16943..17014	Sentido
Tc00.1047053508043.23	ARG	ACG	7578	4377..4448	Antisentido
Tc00.1047053511191.29	ARG	ACG	8619	3869..3940	Sentido
Tc00.1047053506599.6	ARG	CCG	7016	1327..1398	Antisentido
Tc00.1047053508873.405	ARG	CCG	7909	90566..90637	Sentido
Tc00.1047053475029.10	ARG	CCT	4041	319..391	Antisentido
Tc00.1047053509105.118	ARG	CCT	8001	29176..29248	Sentido
Tc00.1047053504427.239	ARG	TCG	6288	53295..53366	Sentido
Tc00.1047053508043.19	ARG	TCG	7578	3714..3785	Sentido
Tc00.1047053475029.30	ARG	TCT	4041	593..665	Antisentido
Tc00.1047053509105.114	ARG	TCT	8001	28902..28974	Sentido
Tc00.1047053475029.50	ASN	GTT	4041	869..941	Antisentido
Tc00.1047053503865.4	ASN	GTT	4791	1555..1627	Antisentido
Tc00.1047053508693.110	ASN	GTT	7832	36684..36756	Antisentido
Tc00.1047053509105.112	ASN	GTT	8001	28627..28699	Sentido
Tc00.1047053510057.60	ASP	GTC	5833	7533..7604	Sentido
Tc00.1047053508909.110	ASP	GTC	7924	35535..35606	Antisentido
Tc00.1047053460125.20	CYS	GCA	3403	942..1013	Sentido
Tc00.1047053510431.55	CYS	GCA	8387	11393..11464	Antisentido
Tc00.1047053511067.17	GLN	CTG	6039	6391..6462	Antisentido
Tc00.1047053511909.6	GLN	CTG	6223	920..991	Sentido
Tc00.1047053510721.16	GLN	CTG	8474	2761..2832	Antisentido
Tc00.1047053511903.132	GLN	CTG	8822	21171..21242	Sentido
Tc00.1047053504427.232	GLN	TTG	6288	52287..52358	Sentido
Tc00.1047053508043.12	GLN	TTG	7578	2719..2790	Sentido
Tc00.1047053504055.89	GLU	CTC	4886	15710..15781	Sentido
Tc00.1047053506435.336	GLU	CTC	6943	66424..66495	Sentido
Tc00.1047053508999.180	GLU	CTC	7958	55260..55331	Antisentido
Tc00.1047053510959.8	GLU	CTC	8554	1190..1261	Sentido
Tc00.1047053506459.246	GLU	TTC	6953	65913..65984	Antisentido
Tc00.1047053506773.123	GLY	CCC	7091	31447..31517	Sentido
Tc00.1047053508799.76	GLY	CCC	7875	15674..15744	Antisentido
Tc00.1047053509715.45	GLY	GCC	8179	8152..8222	Antisentido
Tc00.1047053509717.37	GLY	GCC	8180	7839..7909	Antisentido
Tc00.1047053510105.55	GLY	GCC	8307	9742..9812	Antisentido
Tc00.1047053510105.187	GLY	GCC	8307	42326..42396	Antisentido

Tc00.1047053504427.236	GLY	TCC	6288	52880..52951	Sentido
Tc00.1047053508043.16	GLY	TCC	7578	3296..3367	Sentido
Tc00.1047053511241.10	HIS	GTG	6081	13..84	Antisentido
Tc00.1047053506663.10	HIS	GTG	7045	108..178	Antisentido
Tc00.1047053508087.5	HIS	GTG	7594	587..657	Sentido
Tc00.1047053508861.10	HIS	GTG	7903	1218..1288	Sentido
Tc00.1047053511483.70	ILE	AAT	6131	11150..11223	Antisentido
Tc00.1047053511909.3	ILE	AAT	6223	756..829	Antisentido
Tc00.1047053508777.205	ILE	AAT	7869	44297..44370	Antisentido
Tc00.1047053510721.19	ILE	AAT	8474	2922..2995	Sentido
Tc00.1047053504427.231	ILE	TAT	6288	52128..52201	Antisentido
Tc00.1047053508043.11	ILE	TAT	7578	2560..2633	Antisentido
Tc00.1047053506251.94	LEU	AAG	6871	22235..22316	Sentido
Tc00.1047053506773.126	LEU	AAG	7091	31597..31678	Sentido
Tc00.1047053508231.83	LEU	AAG	7648	28598..28679	Antisentido
Tc00.1047053508799.73	LEU	AAG	7875	15512..15593	Antisentido
Tc00.1047053432939.10	LEU	CAA	16613	643..724	Antisentido
Tc00.1047053507257.200	LEU	CAA	7271	33526..33607	Sentido
Tc00.1047053504427.237	LEU	CAG	6288	52987..53068	Antisentido
Tc00.1047053508043.17	LEU	CAG	7578	3408..3489	Antisentido
Tc00.1047053511909.9	LEU	TAA	6223	1054..1135	Sentido
Tc00.1047053510721.13	LEU	TAA	8474	2621..2702	Antisentido
Tc00.1047053504055.83	LEU	TAG	4886	15284..15363	Antisentido
Tc00.1047053506435.309	LEU	TAG	6943	66005..66084	Antisentido
Tc00.1047053504427.234	LYS	GTT	6288	52564..52636	Sentido
Tc00.1047053504427.242	LYS	GTT	6288	53842..53914	Sentido
Tc00.1047053508043.14	LYS	GTT	7578	2998..3070	Sentido
Tc00.1047053508043.22	LYS	GTT	7578	4263..4335	Sentido
Tc00.1047053475029.20	LYS	TTT	4041	459..530	Antisentido
Tc00.1047053509105.116	LYS	TTT	8001	29037..29108	Sentido
Tc00.1047053504055.87	MET	CAT	4886	15554..15626	Antisentido
Tc00.1047053504055.91	MET	CAT	4886	15866..15938	Sentido
Tc00.1047053506435.327	MET	CAT	6943	66270..66342	Antisentido
Tc00.1047053506435.345	MET	CAT	6943	66577..66649	Sentido
Tc00.1047053506251.88	MET	CAT	6871	22100..22171	Sentido
Tc00.1047053508231.92	MET	CAT	7648	28738..28809	Antisentido
Tc00.1047053506459.243	PHE	GAA	6953	65775..65847	Antisentido
Tc00.1047053506599.9	PHE	GAA	7016	1505..1577	Sentido
Tc00.1047053506603.60	PHE	GAA	7017	17105..17177	Antisentido
Tc00.1047053508873.402	PHE	GAA	7909	90386..90458	Antisentido
Tc00.1047053504033.73	PRO	AGG	4875	15588..15659	Antisentido
Tc00.1047053510339.85	PRO	AGG	8366	20687..20758	Sentido
Tc00.1047053503865.16	PRO	CGG	4791	2045..2116	Sentido
Tc00.1047053508693.140	PRO	CGG	7832	37174..37245	Sentido
Tc00.1047053506251.82	PRO	TGG	6871	21952..22023	Antisentido
Tc00.1047053508231.101	PRO	TGG	7648	28887..28958	Sentido
Tc00.1047053506701.5	SER	AGA	7059	839..919	Antisentido
Tc00.1047053509379.5	SER	AGA	8084	555..635	Antisentido
Tc00.1047053432939.20	SER	CGA	16613	779..859	Antisentido
Tc00.1047053507257.190	SER	CGA	7271	33389..33469	Sentido
Tc00.1047053510057.50	SER	GCT	5833	7342..7422	Antisentido

Tc00.1047053508909.120	SER	GCT	7924	35714..35794	Sentido
Tc00.1047053504055.85	SER	TGA	4886	15446..15526	Sentido
Tc00.1047053506435.318	SER	TGA	6943	66167..66247	Sentido
Tc00.1047053509715.50	THR	AGT	8179	8319..8390	Sentido
Tc00.1047053510105.60	THR	AGT	8307	9909..9980	Sentido
Tc00.1047053503865.8	THR	CGT	4791	1711..1782	Antisentido
Tc00.1047053508693.120	THR	CGT	7832	36840..36911	Antisentido
Tc00.1047053504427.238	THR	TGT	6288	53134..53205	Antisentido
Tc00.1047053508043.18	THR	TGT	7578	3555..3626	Antisentido
Tc00.1047053506847.17	TRP	CCA	5402	1082..1153	Antisentido
Tc00.1047053509459.20	TYR	CCA	8105	5171..5242	Sentido
Tc00.1047053504055.93	TYR	GTA	4886	15975..16060	Antisentido
Tc00.1047053506435.354	TYR	GTA	6943	66685..66770	Antisentido
Tc00.1047053504055.95	VAL	AAC	4886	16159..16232	Sentido
Tc00.1047053506435.363	VAL	AAC	6943	66870..66943	Sentido
Tc00.1047053506459.249	VAL	CAC	6953	66076..66148	Sentido
Tc00.1047053504427.233	VAL	TAC	6288	52419..52491	Sentido
Tc00.1047053508043.13	VAL	TAC	7578	2854..2926	Sentido
Sec 1	SEC	TCA	18653	556..643	Sentido
Sec 2	SEC	TCA	11959	770..857	Antisentido
Sec 3	SEC	TCA	14021	87..174	Antisentido
Sec 4	SEC	TCA	28162	15..102	Sentido
Sec 5	SEC	TCA	11097	441..528	Sentido
Sec 6	SEC	TCA	6957	12609..12696	Antisentido
Sec 7	SEC	TCA	6957	9112..9199	Antisentido
Sec 8	SEC	TCA	6536	26013..26100	Antisentido

APENDICE II. Señales de terminación de la transcripción de genes de RNAt en Trityps.

Las secuencias comienzan inmediatamente después de último nucleótido de los RNAt. Los *clusters* de Ts se muestran en negritas. Aquellos genes que presentan una segunda secuencia (*back-up*) de 4 o más Ts se resaltan en sombreado.

A) *Leishmania major*

Gen	Señal de terminación	Longitud de la secuencia de T	Longitud de la secuencia <i>backup</i>	Distancia entre la secuencia de Ts y <i>backup</i>
LmjF03.TRNALYS.01	CTTTTCCCCCTTGGCTGCGCC	4	-	-
LmjF05.TRNAARG.01	TTTTTTCGTCAGCTCAGAAA	6	-	-
LmjF07.TRNAARG.01	TTTTTCCGGCCCCACTCTCGGGAAC	5	-	-
LmjF09.TRNAARG.01	CATTTTCCAGAGGCATCATG	4	-	-
LmjF09.TRNAGLU.01	CTTTTTCCTTTCCCTGGACAG	5	-	-
LmjF09.TRNAHIS.01	CACTTTTTTCTTCTGTTCTTT	6	-	-
LmjF09.TRNAHIS.02	CACTTTTTTTTTGCCAGCCGG	8	-	-
LmjF09.TRNALEU.01	CACTTTTTCTCCACCTGA	5	-	-
LmjF09.TRNAMET.01	TACTTTTTGCGCTCCTTGTA	5	-	-
LmjF09.TRNAPHE.01	TACCTTTTCTGTTCTTCGAA	4	-	-
LmjF09.TRNAVAL.01	ACTTTTTCCCCGGCCGAAC	5	-	-
LmjF09.TRNAVAL.02	ACTTTTTCTTTCCCTGGACAG	5	-	-
LmjF10.TRNAASN.01	TTTTTCCGGCGCGGGTCATGCTTTTGA AAG	5	4	16
LmjF10.TRNAGLY.01	AAGTTTTTTCGCCACGAGCC	6	-	-
LmjF10.TRNALYS.01	CTTTTTACAAAAAGCCAAC	5	-	-
LmjF11.TRNAALA.01	CATTTTCGTCGCAGCGGAGAT	4	-	-
LmjF11.TRNAALA.02	CATTTTCCCCTAATTCAGT	5	-	-
LmjF11.TRNAARG.01	ACTTTTCTGGGCCGTCAAA	4	-	-
LmjF11.TRNAGLY.01	CACTTTTCCCGCAAAGGGTTC	4	-	-
LmjF11.TRNALEU.01	CTTTTCTGCCACTGTCTCT	5	-	-
LmjF11.TRNALEU.02	CTTTTCTGCCACTGTCTCTG	5	-	-
LmjF11.TRNAMET.01	AACTTTTTCTTACGTCCTC	5	-	-
LmjF15.TRNAASP.01	CTTTTCCTATGGTGCACGGCG	4	-	-
LmjF15.TRNAGLU.01	CTTTTCCGCCCGGGGCGCC	4	-	-
LmjF16.TRNAGLN.01	ACTTTTTCCCCGAGCCCCTC	5	-	-
LmjF17.TRNAALA.01	ACTTTTGTGTGGAGCCGGCG	4	-	-
LmjF17.TRNAASP.01	CTTTTGGCCGAAGCCCGG	5	-	-
LmjF17.TRNASER.01	CTTTTCCGAAACCTGTGGCACTACTC CGAAGATTTT	4	4	28
LmjF21.TRNALYS.02	CTTTTTCGTAGCAGCGCGCC	5	-	-
LmjF21.TRNAPRO.01	TTTTCCGCCACCCTTTCCCGCGGTGC CATTTTCTGC	4	4	25
LmjF21.TRNASER.01	TACTTTTTCCGCCATCACGC	5	-	-
LmjF21.TRNAVAL.01	CTTTTCTTGCTTGCCTTTATT	4	-	-
LmjF23.TRNAARG.01	TTTTCTTTGCTTTGCCCC	4	4	1
LmjF23.TRNAARG.02	CATTTTTCTATCGTTAGGCA	5	-	-

LmjF23.TRNAGLN.01	GTTTTTGGGCTGGAGCAGCT	5	-	-
LmjF23.TRNAGLY.01	CTGTTTTCCCAGTGCCCCGC	4	-	-
LmjF23.TRNAILE.01	GCCGCTCTTTTTTCGCGAGCCC	5	-	-
LmjF23.TRNALEU.01	ACCTTTTCTATAAGC	4	-	-
LmjF23.TRNALYS.01	ATTTTCCCTTTTGCAGGCG	4	4	3
LmjF23.TRNATHR.01	TTTTTGGTTTTTCGGGCCAT	5	4	2
LmjF23.TRNATRP.01	TTTTTTTTTGTCTGCGGTGT	9	-	-
LmjF23.TRNAVAL.01	TTTTTCTCTGCGGAGCGAAT	5	-	-
LmjF24.TRNAASP.01	GTTTTTCCTTGCCCAACACC	5	-	-
LmjF24.TRNAGLN.01	ACTTTTTCCCCGAGTACAAAC	5	-	-
LmjF24.TRNAILE.01	CTTTTTCCAACCCCCAGAATAGGGT GTCCGGCTTTTTTT	5	7	28
LmjF24.TRNAILE.02	CTTTTTCTAACCCCCAGAATAGGGG TGCCGGCTTTTTTTTTT	5	10	29
LmjF24.TRNALEU.01	AAATGTTTTCAGCAT	4	-	-
LmjF24.TRNAPRO.01	ATTTTTTTTGTGTGTGGAGA	8	-	-
LmjF29.TRNALEU.01	ATTTTTTCCCCGACCACAAG	6	-	-
LmjF29.TRNASER.01	GTTTTCAACTCACTCCT	5	-	-
LmjF30.TRNATHR.01	TGCCTTTTTCTGGCGCCAGG	5	-	-
LmjF31.TRNAALA.01	ACTTTTTCACCCGACAAGTGA	5	-	-
LmjF31.TRNAGLU.01	AATTTTTTCCAGGCCGC	6	-	-
LmjF31.TRNAGLY.01	AACGTTTTCCCTAGCCGCC	4	-	-
LmjF31.TRNAPHE.01	TACTTTATTTTCCCAGCCC	4	-	-
LmjF31.TRNASER.01	CTTTTCGAGAAGCAAGCCC	4	-	-
LmjF33.TRNAALA.01	CTTTTTTCTGTTCGGGGCC	6	-	-
LmjF33.TRNAARG.01	CGTTTTTGGAGAGGCTTTTTCAA	5	5	9
LmjF33.TRNAARG.02	CTTTTGATAGGTGCAGCAAG	4	-	-
LmjF34.TRNAASN.01	CTTTTCGCCGACG	4	-	-
LmjF34.TRNAASN.02	CTTTTCTTTACACACGGTGAA	4	-	-
LmjF34.TRNAILE.01	CTTTTTTCAACACGATGAAACATATT TTTGAGG	7	5	17
LmjF34.TRNALEU.01	ACCCTTTTCCAGGGTTCGGGTTTTG	4	4	12
LmjF34.TRNA MET.01	TCCCTTTTTTCAAAACCCGAC	6	-	-
LmjF34.TRNAPRO.01	TTTTCCCTCTAGGCCTCGGG	4	-	-
LmjF34.TRNASER.01	TGTTTTTTCGCCCCGCCCA	6	-	-
LmjF34.TRNATHR.01	GTCTCCTTTTTCAGCCTCTGG	4	-	-
LmjF34.TRNATYR.01	AGTTTTTCCTTCCTCTCGAAGGAC	4	-	-
LmjF34.TRNAVAL.01	CTTTTTTACCCAGACGGTAG	6	-	-
LmjF36.TRNACYS.01	TCTTTTCCCCTGGATTGGGGCAG	4	-	-
LmjF36.TRNAGLN.01	TCTTTTTCCAAGCCTCACC	5	-	-
LmjF36.TRNAGLY.01	GGCTTTTCTCGCCTAAGCCAG	4	-	-
LmjF36.TRNAGLY.02	AAGTTTTTTGACACACTGCTGGTG TTTTATGG	6	4	15
LmjF36.TRNALEU.01	CTTTTCTTCCACCCAAA	4	-	-
LmjF36.TRNA MET.01	TACGTTTTTCGCCGAGAAAAC	4	-	-
LmjF36.TRNAPRO.01	TTTTTCCTCCAAGCGGA	5	-	-
LmjF36.TRNAPRO.02	CTTTTTCCGACGGAGTAAAGA	5	-	-
LmjF36.TRNATHR.01	CTTTTTTCCCCACGTACCCT	4	-	-
LmjF36.TRNATHR.02	CTTTTTTCCCCACGTACCCTC	4	-	-
LmjF36.TRNATHR.03	AACTTTTTTGCATCGAAAAG	6	-	-

LmjF36.TRNATYR.01	CACCTTTTTCCCAGCCGCC	5	-	-
LmjF36.TRNATYR.02	CACCTTTTTCCCAGCCGCC	5	-	-
Chr06.SeC	TGCCATTTTCCTTTTCGCTCCGA	4	4	2

B) *Trypanosoma brucei*

Gen	Señal de terminación	Longitud de la secuencia de T	Longitud de la secuencia backup	Distancia entre la secuencia de Ts y backup
Tb04_trNA_Glu_1	ACTTTTTTAGAGGT	7	-	-
Tb04_trNA_Leu_1	GTCGTTTTCGTGTAAATCA	4	-	-
Tb04_trNA_Met_1	TATTTTACTGCGGAGTAAG	5	-	-
Tb04_trNA_Met_2	CTTTTTGATTTCCACCTGCGAG TTTTTC	6	5	16
Tb04_trNA_Pro_1	GCTTTTCCGCATCTAGTG	5	-	-
Tb04_trNA_Ser_1	GGTTTTCTCCTGCTGAAAAA	4	-	-
Tb04_trNA_Thr_1	CGTTTTTTCGCCGACCAAGT	6	-	-
Tb04_trNA_Tyr_1	CTTTTTCACGCCACAAACGAGT	5	-	-
Tb04_trNA_Val_1	GGTTTTCGTCATTTCTCCGTTTCCC TCAATTTTGCTGCCGT	4	4	23
Tb04_trNA_Val_2	GGTTTTCGTCATTTCTCCGTT TCCCTCAATTTTGC	4	4	23
Tb04_trNA_Val_3	AGCCTTTTTATATTCCCGTT	5	-	-
Tb05_trNA_Arg_1	CTGTTTTTTTCCCCTCGT	7	-	-
Tb05_trNA_Glu_1	GCCTTTTTGGAGACCAACGT	6	-	-
Tb07_trNA_Ala_1	CGTTTTCGTCAACCGC	4	-	-
Tb07_trNA_Ala_2	CGTTTTCGTCAACCGCATCC	4	-	-
Tb07_trNA_Ala_3	CGTTTCCATCCACTCAATTC TCACACTTTTTC	4	4	22
Tb07_trNA_Asp_1	GTTTTTTCGTCTGTGTACA	6	-	-
Tb07_trNA_Asp_2	GTTTTTTCGTCTGTGTACAC	6	-	-
Tb07_trNA_Gly_1	CGTTTTTCTCCCATCTGTGCG TTTTGCATTGG	5	4	14
Tb07_trNA_Lys_1	CTTTTCCCCTTCGGTCCAGG	4	-	-
Tb07_trNA_Ser_1	GTTTTTCCCCTTTGAGGTCA TCGTGTTTTATAG	5	4	20
Tb07_trNA_Ser_2	GTTTTTCCCCTTTGAGGTCACTCG TGTTTTATA	5	4	20
Tb08_trNA_Arg_1	CACTTTTCTCAGTCACATTGGT TCGATAGCGCC	4	-	-
Tb08_trNA_Arg_2	TTTTTACCTAAGAGAAGGACGG GGCAGACGGTTTTGACA	5	4	27
Tb08_trNA_Arg_3	CTTTTCCCTCTTCAGGCGTC	4	-	-
Tb08_trNA_Asn_1	GCTTTTTTCCATAGCGAGTG	7	-	-
Tb08_trNA_Cys_1	TAGTTTTCCGGAGGAGGGAA	4	-	-
Tb08_trNA_Cys_2	TAGTTTTCCGGAGGAGGGAAAG	4	-	-
Tb08_trNA_Gln_1	AATTTTCGTTTTTTCAATTC	4	7	2
Tb08_trNA_Gln_2	CTTTTGGACACATAACGCCACAGT	5	-	-
Tb08_trNA_Gln_3	CTTTTGGACACATAACGCCACAGT	5	-	-
Tb08_trNA_Gly_1	CATTTTTTCGTTTTATCAT	6	5	2
Tb08_trNA_Ile_1	TTCGTTTTCTTTGCGGAG	4	-	-
Tb08_trNA_Ile_2	CGTTTTTCAGTCCGGGAGACT	5	-	-

Tb08_tRNA_Ile_3	CTTTTCTCTACCTATCAGTT CGTTGTTTTTAC	5	4	20
Tb08_tRNA_Leu_1	ATATTTTCGTCCGATCATATC	4	-	-
Tb08_tRNA_Leu_2	GTTTTTGCCTTTAG	6	-	-
Tb08_tRNA_Lys_1	CTTTTCTCCGATTTCCTGA	4	-	-
Tb08_tRNA_Lys_2	CTCTTTTTTTTTGACTTCCC	10	-	-
Tb08_tRNA_Thr_1	TTGTTTTCGGCCTATTGAC	4	-	-
Tb08_tRNA_Trp_1	TTTTCCCCAACACACT	5	-	-
Tb08_tRNA_Val_1	GCGTTTTCCAGTTGCGGCTG	4	-	-
Tb09_tRNA_Gly_1	TCGTTTTTCTGCCGATGCAGA ACCCACCATTTTTTTTA	5	9	22
Tb10_tRNA_Ala_1	CTGTTTTGCAACAGCT	5	-	-
Tb10_tRNA_Arg_1	CGTTTTCGTG	4	-	-
Tb10_tRNA_Arg_2	GTTTCCGGTGCAAAAAC	4	-	-
Tb10_tRNA_Asn_1	TCTTTTGCTGC	5	-	-
Tb10_tRNA_Gly_1	GACTTTCACGGCTC	4	-	-
Tb10_tRNA_Leu_1	ACTTTTGCGAACCGGG ACCTTATTTTC	5	4	18
Tb10_tRNA_Lys_1	CTTTTCTACACTGCCGCTTAA	4	-	-
Tb10_tRNA_Met_1	AGTTTTACATCCAC CCTTTTTCTCC	4	6	11
Tb10_tRNA_Pro_1	TTTTCTGGGCGCCCAATTTCC	4	-	-
Tb10_tRNA_Pro_2	TCTTTTCGATTTCCTTGTCAGG	4	-	-
Tb10_tRNA_Thr_1	CGTTTTTACCACCTTCAAA	5	-	-
Tb11_tRNA_Ala_1	TCTTTTTGCTTACTTGTGC	6	-	-
Tb11_tRNA_Arg_1	GTGTTTTCGTCTACCCTGTG	4	-	-
Tb11_tRNA_Glu_1	GTTTTTCGCAGTTCAATGCG	5	-	-
Tb11_tRNA_Gly_1	TCTTTTTCGCGTCACTCCTTTTTCG	5	4	13
Tb11_tRNA_His_1	CACGTTTTCTTATGATTTTCG	4	-	-
Tb11_tRNA_Phe_1	AAGTTTTTGCACCTCCTCCCA	5	-	-
Tb11_tRNA_Phe_2	AAGTTTTTGCACCTCCTCCC	5	-	-
Tb927.3.4393 Chr3-Ser	TTTTTTTCGTCAATTCCCCGTGTTCCG CTGCTTTTG	6	4	24
Tb927.3.4396 Chr3-Leu	CTTTTTTCGCACCCAAATCGGCTGATT GTTTTCA	5	4	21
Tb927.3.4399 Chr3-Ser	GTTTTTGGCATTGCAT	5	-	-
Chr09.Sec.1	TTTTTAATGTTGTTGTTGTTTTGT	5	4	13
Chr09.Sec.2	TTTTTAATGTTGTTGTTGTTTTGT	5	4	13

C) *Trypanosoma cruzi*

Gene	Señal de terminación	Longitud de la secuencia de T	Longitud de la secuencia backup	Distancia entre la secuencia de Ts y backup	
Tc00.1047053510057.40	ALA	ACTTTTTTCCGATTTTTTTTTTTTTG	6	13	4
Tc00.1047053508909.130	ALA	ACTTTTTTTCGATTTGTTTTTTTTTCGC	6	8	7
Tc00.1047053506599.3	ALA	TTTTTCCTTTTTTTTTTGAAAA	5	8	2
Tc00.1047053508873.408	ALA	CTTTTTCTCCATTTTTTTGATG	5	6	6
Tc00.1047053475029.40	ALA	CGTTTTTTTACACTAAACATTTTTT	6	6	10
Alanina (no anotado)	ALA	TGTTTTTGCACGAAAAATTTTTTTTCG	5	7	12
Tc00.1047053504427.243	ARG	GTTTTTTTTCAAAAATTTTTC	8	5	5
Tc00.1047053506619.59	ARG	AATTTTTTTGCCTATTGCGG	7	-	-

Tc00.1047053508043.23	ARG	ACTTTTAGAAGATGTTTTTC	4	5	8
Tc00.1047053511191.29	ARG	CGGATGTTTTTTTTTAGTTATTTTC	10	4	5
Tc00.1047053506599.6	ARG	GTTTTTTCTCATTGTTTTAAACATTG	7	4	7
Tc00.1047053508873.405	ARG	TTTTTTTTTATATGTTTTTCTC	10	6	6
Tc00.1047053475029.10	ARG	ATTTTTTTTTTTTTTTTCA	16	-	-
Tc00.1047053509105.118	ARG	ATTTTTTTTTTTTTTTTCATTC	16	-	-
Tc00.1047053504427.239	ARG	TTTTTTTTTCCCCTATTTTTTGCTGGCA	9	6	6
Tc00.1047053508043.19	ARG	TTTCTTTCCCCTATTTTTTTACTAG	4	8	11
Tc00.1047053475029.30	ARG	TATATTTTTTTCCCTTTGTGC	7	-	-
Tc00.1047053509105.114	ARG	TATATTTTTTTCCCTTTGTGCA	7	-	-
Tc00.1047053475029.50	ASN	TTTTTTCGTGCTTCGAAC	6	-	-
Tc00.1047053503865.4	ASN	TTTTTTTCGCACCATTTTTTCATTCTC	7	6	7
Tc00.1047053508693.110	ASN	ATTTTTTTTCGCTCCA TTTTTTTTTTTTTTTTATTTCTAGGAGC	7	15	7
Tc00.1047053509105.112	ASN	TTTTTTTGTGCTTCGAAAAAA ATTTTTTCGTGTC	7	6	16
Tc00.1047053510057.60	ASP	TTTTTTTCACCTCGTGAAAAAT	7	-	-
Tc00.1047053508909.110	ASP	ATGCTTTTGTGGTTGTTCGAG	4	-	-
Tc00.1047053460125.20	CYS	TACTTTTTCAAAAATTATTTTA	6	4	9
Tc00.1047053510431.55	CYS	TACTTTTCAAAAAAAAAAATTTA CTTTATTTT	4	4	21
Tc00.1047053511067.17	GLN	ACTTTTTTCTCTTATTTATT	6	-	-
Tc00.1047053511909.6	GLN	ACTTTTTTCTCTATGTTTTTTGAGG	5	7	6
Tc00.1047053510721.16	GLN	ATTTTTTATTATATTTTTTA	6	6	6
Tc00.1047053511903.132	GLN	ACTTTTTTCTCTTATTTAT	5	-	-
Tc00.1047053504427.232	GLN	CTTTTTTTTTTGAAAAATGT	10	-	-
Tc00.1047053508043.12	GLN	CGTTTTTTTGTGTTGAGGAAA	7	-	-
Tc00.1047053504055.89	GLU	TCTTTTTCCCTTGCACGGTG	5	-	-
Tc00.1047053506435.336	GLU	TTTTTCCTGGACAGC	5	-	-
Tc00.1047053508999.180	GLU	TTTTTTCCAAAAAAAAAATT	6	-	-
Tc00.1047053510959.8	GLU	ACTTTTTTCCAAAAAAAAAAT	6	-	-
Tc00.1047053506459.246	GLU	CTTTTTTGGTCTCAAAAC	6	-	-
Tc00.1047053506773.123	GLY	CATTTTTTGGTTTTATCAAAT	6	4	2
Tc00.1047053508799.76	GLY	CATTTTTTAGTTTTATCAC	6	4	2
Tc00.1047053509715.45	GLY	GACTTTTTTTGCTCT	8	-	-
Tc00.1047053509717.37	GLY	GACTTTTTGCTCTCCAAAAAGG	5	-	-
Tc00.1047053510105.55	GLY	AACTTTTTTTTGCCTTTC	9	-	-
Tc00.1047053510105.187	GLY	GACTTTTTTTTTTGCCTTT	11	-	-
Tc00.1047053504427.236	GLY	AACTTTTTCCATTTTTTGC	5	6	4
Tc00.1047053508043.16	GLY	ACTTTTTCTCCTATTCT	4	-	-
Tc00.1047053511241.10	HIS	TAATTTTTCGGA	5	-	-
Tc00.1047053506663.10	HIS	TACTTTTTCGGAGTTTTTC	5	5	5
Tc00.1047053508087.5	HIS	TACTTTTTCGGAGTTTTTCCC	5	5	5
Tc00.1047053508861.10	HIS	TAATTTTTCGGAGTTTTTCTCCGAGCG AA	5	5	5
Tc00.1047053511483.70	ILE	CTTTTTCTCCAGAAATAAAT	5	-	-
Tc00.1047053511909.3	ILE	CTTTTTGTCCAACTTCGGC	5	-	-
Tc00.1047053508777.205	ILE	CTTTTTCTCTATGAAATAAA	5	-	-
Tc00.1047053510721.19	ILE	CTTTTTTGTCCAAATTCTG	6	-	-
Tc00.1047053504427.231	ILE	AAGTTTTATTTTAGACGCTGT	4	4	1
Tc00.1047053508043.11	ILE	AAGTTTTATTTCTAGACGCT	4	-	-
Tc00.1047053506251.94	LEU	ATTTTTTCCCCTTTGGGCTT	6	-	-
Tc00.1047053506773.126	LEU	ATTTTTTCCCCTTGCTTTTTTTTTATGC	6	9	8

		ATG			
Tc00.1047053508231.83	LEU	ATTTTTTCCCTTTTTTGGCT	6	6	3
Tc00.1047053508799.73	LEU	ATTTTTTCCCTTTTTGC	6	5	3
Tc00.1047053432939.10	LEU	CTTTTTTCGCAATCGAGGGT	6	-	-
Tc00.1047053507257.200	LEU	CTTTTTTCGCAATCGACAGCG	6	-	-
Tc00.1047053504427.237	LEU	ATTTTTTTCCTCTCTGCAAA	7	-	-
Tc00.1047053508043.17	LEU	ACTTTTTTCCTCCTTCGCA	5	-	-
Tc00.1047053511909.9	LEU	ATGTTTTTCCGCGGAAGCCAT	5	-	-
Tc00.1047053510721.13	LEU	ATTTTTTTTTTTTTCGCAGAAA	12	-	-
Tc00.1047053504055.83	LEU	GGAAGTTTTTCGATTTTTTTG CTTTTTTTCGCGCAA	4	7	3
Tc00.1047053506435.309	LEU	AGGAAGTTTTCGAATC	4	-	-
Tc00.1047053504427.234	LYS	TTTTTTTCGTTGCTTGTGTTTTTTTTT CT	7	10	11
Tc00.1047053504427.242	LYS	TTTTTTTAAAAGAGAAAAATTTTGA	7	4	12
Tc00.1047053508043.14	LYS	TTTTTTAATGTATTTTTCTC	6	5	6
Tc00.1047053508043.22	LYS	TTTTTTTTTTTCTTTAAAAA	11	-	-
Tc00.1047053475029.20	LYS	CTTTTTTCCAACCTTTCAT	7	-	-
Tc00.1047053509105.116	LYS	CTTTTTTCCAACCTTTCATTTTTGA	7	5	10
Tc00.1047053504055.87	MET	TTTTTTCCTTTTTCTTGGGA	6	5	2
Tc00.1047053504055.91	MET	TTTTTTTTCTCCTGC TTTTTTTTTTTGCAAA	8	11	7
Tc00.1047053506435.327	MET	TTTTTTTTCTTGGGAAAAA	9	-	-
Tc00.1047053506435.345	MET	TTTTTTTCCTGCGTTTTTTTG	7	7	6
Tc00.1047053506251.88	MET	AATTTTTTGTTTTTTTTCGT	7	8	1
Tc00.1047053508231.92	MET	AATTTTTTGTTTTTTTCTTT	6	7	1
Tc00.1047053506459.243	PHE	TAAGTTTTTCCCAAACTTTTTTTC	5	6	7
Tc00.1047053506599.9	PHE	TAGTTTTTCCCAAATTTTTTTTAA	5	8	7
Tc00.1047053506603.60	PHE	TAGTTTTTCCCAAACTTTTTTAA	5	6	8
Tc00.1047053508873.402	PHE	TAAGTTTTTCCCAAACTTTTTTA	5	6	7
Tc00.1047053504033.73	PRO	TTTTTTTCACCAGAAAGATG	7	-	-
Tc00.1047053510339.85	PRO	TTTTTTTCCAGAAAGACGCG	7	-	-
Tc00.1047053503865.16	PRO	TGTTTTTTTTTCTTTTTAAAT	9	5	1
Tc00.1047053508693.140	PRO	TGTTTTTCCCAAAGTCA	6	-	-
Tc00.1047053506251.82	PRO	TTTCTTTTAAACCGAAAAA	4	4	1
Tc00.1047053508231.101	PRO	TTTTTTTCGGGCCGAAAAAATT	7	-	-
Tc00.1047053506701.5	SER	ATTTTTTCCCTTCCAGCCATTGAGC	6	-	-
Tc00.1047053509379.5	SER	ACTTTTTTTCACCTTCCAAC	7	-	-
Tc00.1047053432939.20	SER	GTTTTTTCCTTCCAGCCAT	6	-	-
Tc00.1047053507257.190	SER	GTTTTTTTTCTTTCCAACCA	8	-	-
Tc00.1047053510057.50	SER	ATGTTTTTCTCATATCAAAA	5	-	-
Tc00.1047053508909.120	SER	ATGTTTTTCCCAAATTA AAAA	5	-	-
Tc00.1047053504055.85	SER	GTTTTTTTTCCAAGAAAAAG	8	-	-
Tc00.1047053506435.318	SER	GTTTTTTCCTAAGAAAAAAA	6	-	-
Tc00.1047053509715.50	THR	CTTTTTTCCGCGCCGATTC	7	-	-
Tc00.1047053510105.60	THR	CTTTTTTCTGGGCCGAAT	7	-	-
Tc00.1047053503865.8	THR	GCTTTTCTCTTTTTTTTTTTT	5	12	3
Tc00.1047053508693.120	THR	GCTTTTCTCTTTTTTTTTTTT	5	12	3
Tc00.1047053504427.238	THR	TTTTCCTTCTCAGCAATTGA	4	-	-
Tc00.1047053508043.18	THR	TTTTCCTTCTCAGCAATTCG	4	-	-
Tc00.1047053506847.17	TRP	AAGTTTTTTTGTTCCTCCGAT	7	-	-
Tc00.1047053509459.20	TYR	GATTTTTTCTGTTCCCGCG ATTTCCATTTTTTTTTTTTC	6	13	18

Tc00.1047053504055.93	TYR	ATCTTTTTGCAAAAAAAAAA	5	-	-
Tc00.1047053506435.354	TYR	ACCTTTTCGCAAAACAAAAA	4	-	-
Tc00.1047053504055.95	VAL	CTTTTCTCCCATACAGCA	4	-	-
Tc00.1047053506435.363	VAL	CTTTTTCATCCAGC	6	-	-
Tc00.1047053506459.249	VAL	GTTTTTTCAGCATCGACGCTT	6	-	-
Tc00.1047053504427.233	VAL	TCGTTTTCCAAAAGTTTA	4	4	7
Tc00.1047053508043.13	VAL	TCGTTTTCCAAGAGCTTCCAA	4	-	-
Contig 18653	Sec	TTTTTTTTTTTCGGCGAACTGCAGTC CGTCTTTTTTATTCTTTGTCCGTTGTGT	11	6	19
Contig 11959	Sec	TTTTTTTTTTTCGGCGAACTGCAGTC CGTCTTTTTTATTCTTTGTCCG	11	6	19
Contig 14021	Sec	TTTTTTTTTTTCGGCGAAAACCTGCAGTCT GTCTTTTTAATTCTTTGTCCGT	9	5	22
Contig 28162	Sec	TTTTTTTTTTTCGGCGAACTGCAGTCCG TCTTTTTTATTCTTTGTCCGTT	11	6	19
Contig 11097	Sec	TTTTTTTTTTTCGGCGAACTGCAGTCCG TCTTTTTTATTCTTTGTCCGTT	11	6	19
Contig 6957 (a)	Sec	TTTTTTTTTCGGCGAACTGCAGTCTGT CTTTTAAATTCTTTGTCCGTTG	9	5	20
Contig 6957 (b)	Sec	TTTTTTTTTCGGCGAACTGCAGTCTGTC TTTTTAAATTCTTTGTCCGTTGT	8	5	20
Contig 6536	Sec	TTTTTTTTTTTCGGCGAACTGCAGTCCG TCTTTTTTATTCTTTGTCCGTT	11	6	19